The structure and scale of the universe

Hoyle, Fiona

The Structure and Scale of the Universe

by Fiona Hoyle

A Thesis submitted to the University of Durham
in accordance with the regulations for
admittance to the Degree of Doctor of Philosophy.

Department of Physics
University of Durham

July 2000

Abstract

The Structure and Scale of the Universe

We quantify the structure and scale of the Universe using redshift surveys of galaxies and QSOs and observations of Galactic open star clusters.

We obtain the galaxy power spectrum from the Durham/UKST Galaxy Redshift Survey. By comparing the shape of the observed power spectrum to the APM real space power spectrum, we quantify the size of the redshift space distortions and find $\beta = \Omega^{0.6}/b$=0.60±0.35. We also apply counts-in-cells analysis to the Durham/UKST and Stromlo-APM Surveys and measure the skewness directly out to $20h^{-1}$Mpc. We find that the skewness measured from CDM models can only be reconciled with that of galaxies if bias is non-linear.

We make predictions for the clustering in the 2dF QSO Survey by constructing mock catalogues from the *Hubble Volume* N-body simulation, with geometry, selection function and clustering matching those expected in the completed Survey. We predict that the correlation function will be reliably measured out to $\sim 1,000h^{-1}$Mpc and the power spectrum out to $500h^{-1}$Mpc. We measure the power spectrum from the 2dF QSOs observed by January 2000 and find it has a shape of $\Gamma \sim 0.1$. We also find little evolution in the clustering amplitude as a function of redshift. We obtain constraints on the cosmological parameters $\Omega_{\rm m}$ and β by combining results from modeling geometric distortions introduced into the clustering pattern due to inconsistent cosmological assumptions and results from the QSO-mass bias.

Finally, we consider the scale of the Universe. We check the calibration of the Cepheid Period-Luminosity relation using U, B, V and K band imaging of Galactic Open Clusters containing Cepheids and measure the distance modulus to the LMC to be 18.51 ±0.10. However, we find anomalous colour-colour diagrams for two clusters and suggest that the effects of metallicity may be greater than previously considered.

Contents

List of Figures

List of Tables

Preface

The Structure and Scale of the Universe

The work described in this thesis was undertaken between October 1997 and July 2000 whilst the author was a research student under the supervision of Dr Tom Shanks in the Department of Physics at the University of Durham. This work has not been submitted for any other degree at the University of Durham or at any other University.

The work presented in this Thesis is my own, unless otherwise explicitly stated. I played a leading role in all the work presented in this thesis. Chapter 3 was done in collaboration with Carlton Baugh and Tom Shanks and Chapter 4 in collaboration with Carlton Baugh and Istvan Szapudi. Chapter 8 was done in collaboration with Tom Shanks and Nial Tanvir (University of Hertfordshire).

A number of the results have appeared in the following papers

- Hoyle F., Baugh C. M., Shanks T. & Ratcliffe A., 1999, MNRAS, 309, 659
- Hoyle F., Szapudi I. & Baugh C. M., MNRAS accepted
- Hoyle F., Shanks T. & Tanvir N. R., MNRAS submitted
- Hoyle F., Shanks T. & Tanvir N. R., 1999, in 'Harmonising the Cosmic Distance Scale in a Post-Hipparcos Era', eds Egret D & Heck A., ASP Conference series, 167, 198

Acknowledgements

It is at this part of the thesis where I am given chance to thank everyone who has helped me over the last three years.

Firstly I must thank my supervisor, Tom Shanks. Tom certainly ensures that the PhD years are lively and stimulating and an experience never to be forgotten. Just don't let him start talking about cows or conspiracy theories! Carlton Baugh and Phil Outram deserve special mention for helping me get to grips with various aspects of my work and for reading early drafts of this Thesis. Thanks also to other collaborators – Istvan Szapudi, Nial Tanvir and Scott Croom. I'd also like to thank my undergraduate supervisor, Judith Perry, for having belief in me which lead to me starting the PhD in the first place.

Office mates and fellow post grads/post docs in Durham who have kept me sane (some would disagree!) deserve lots of thanks: Mikey and Carlos for those early dinner parties, thoughts of tequila still haunt me. Russell and Simon for nights out in the pub, beer will never cost £1.24 a pint again. Katherine, Harald, Claire, Eric, Dave, Kev, Richard and Amaya deserve lots of praise for putting up with me singing in the office! Peder needs thanking for keeping the Durham Astro Tennis Competition going, although that may just be because he knows he can win it! These are just a few of the people who have helped make the three years at Durham fly by. Thanks go to my old school and university friends too. Ursula and Comp deserve a mention as they have managed to make it up to see the splendours of Durham.

My family have given me immense amounts of support over the years, particularly while I have been a student. Thanks go to Mum for keeping me supplied with black cardigans and to Dad for his tales of the garden in Shepley. Yes I am really pleased you have a newt in the pond! Thanks go to my brother Alex for sending me lots of emails from his trip round Oz. It was great to relive my memories whilst being up to my neck in thesis drafts!

No surprises, a final, huge thanks goes to Andrew. You're the best pony head I've ever met!

This thesis is dedicated to the treehouse

Chapter 1 | *Introduction*

For the last few decades, the key questions in cosmology have remained essentially the same. We still wish to know when the Universe began, what its ultimate fate will be, what it is made of, what form the structure in the Universe takes and how the structure was formed.

One hundred years ago, observations of what we now know as galaxies external to our own were interpreted as being nebulae within the Milky Way which would eventually collapse to form planetary systems. The size of the Galaxy was also unknown. Estimates placed our Sun at only 650 pc from the centre of the Galaxy, as no correction had been made for dust that lies along the line of sight. This was known as the Kapteyn Universe. It was also thought that the Universe was static. (For more details see, for example, Shu 1982)

During the 1920's, observations of the local Universe began to change this picture. Hubble discovered the expansion of the Universe, through the Cepheid Period-Luminosity relation, and it was realised that the objects that were once thought of as being within our Galaxy, actually lay at great distances and were different galactic systems. It was also around this time that Einstein's General Theory of Relativity was starting to be accepted as the theory which describes gravity on cosmological scales.

Today, most astronomers would agree on the general properties of the Universe, such as the existence of galaxies external to our own, the fact that the Universe is expanding and that the age of the Universe is around 15±5Gyr. Although the questions posed at the start of this section have yet to be fully answered, progress has been made during the 20^{th} century towards a better understanding of the nature of the Universe.

1.1 The Standard Cosmological Model

The standard model assumes the *Cosmological Principle*. This principle states that the Universe is spatially homogeneous and isotropic on sufficiently large scales at any particular cosmic time. This assumption, coupled with the dynamical equations of general relativity, has had remarkable success in predicting a wide range of observable proper-

ties. The first of these to be measured was the expansion of the Universe. During the 1920's, Hubble measured the recession velocities of 45 nearby galaxies and published the result that recession velocity was proportional to the galaxy's distance from the observer (Hubble 1929). The expansion of the Universe was originally predicted by Einstein but he introduced the cosmological constant to keep the Universe static, in line with pre-Hubble thinking. Upon Hubble's discovery of the expanding Universe, Einstein named the cosmological constant his 'biggest blunder', although, in recent times, the cosmological constant has had a revival, although for a different reason. In this Thesis we even describe a way in which it may be detected.

The expansion of the Universe suggests that the Universe was smaller in the past, and therefore denser than it is today. This led to the idea of the Universe starting in a hot 'Big Bang'. In 1965, Penzias and Wilson serendipitously detected the remnant of this hot, dense phase when they found that the Universe possesses a cosmic microwave background radiation (CMBR) (Penzias & Wilson 1965). This has been investigated in more detail by the COBE satellite and is found to be a near perfect black body, radiating at a temperature of 2.726 ± 0.010K (Mather et al. 1994) and uniform to 1 part in 10^5 (Smoot et al. 1992). The other observable prediction of the standard model is the abundance of light elements, such as hydrogen and helium arising from primordial nucleosynthesis. At early times, the model predicts that the Universe was very hot, $\gtrsim 10^9$K, and dense and nuclear reactions could occur, producing the light elements. There is good agreement between the prediction of the standard model and the inferred abundances of hydrogen, helium, deuterium and lithium (Schramm 1991). These three observations provide strong evidence that the framework in which most cosmologists work accurately describes the Universe.

1.2 Structure in the Universe

The Cosmological Principle seems to hold on large scales, as demonstrated by, for example, the high level of uniformity seen in the CMBR. However, on smaller scales, observations show that the matter is not homogeneously distributed but clumped together in the form of stars, galaxies and clusters. This then begs the question 'how did the Universe become so inhomogeneous on small scales when the CMBR is so smooth?' The most widely accepted picture is that the small fluctuations initially grew linearly via gravitational instability. As fluctuations reached overdensities $\sim$1 the growth became non-linear, causing

the fluctuations to grow much more rapidly than linear theory suggests (see Appendix B). Eventually the structure that we see in the Universe today was formed.

In order to understand how the matter fluctuations evolve, we need to know the values of the cosmological parameters, which drive the evolution of the fluctuations, and the distribution of the mass at present day. However, measurements of clustering are made, for example, from observations of galaxies and quasi-stellar objects (QSOs), the latter of which are believed to be the nuclei of active galaxies. These objects may not trace the distribution of the mass in a simple way. In fact, the visible parts of galaxies seem to comprise only a small fraction of the total mass, with most of the material in the Universe being invisible dark matter (Zwicky 1933, Bahcall 1977, Faber & Gallagher 1979). The nature of this dark matter is as yet unknown. Various candidates have been postulated and are discussed further in Chapter 2.

Presently the clustering of galaxies is known over a fairly wide range of scales ($0.1 \lesssim r \lesssim 100h^{-1}$Mpc*) and the next generation of surveys will extend this range. However, direct measurements of the clustering of mass in the Universe are far more difficult. Perhaps the most promising method of measuring the clustering of the mass over a range of scales comes from studies of the amount of absorption seen in the spectrum of high redshift QSOs, known as the Lyman α forest, (Croft et al. 1998). There are many uncertainties in this method and the range of scales over which this can be applied are limited.

1.3 The Scale of the Universe

The recession velocities of galaxies are seen to be proportional to their distance from us and the constant of proportionality is known as the Hubble constant, H_o, where the subscript o refers to the present day. In the past, the constant of proportionality was larger than it is at present day and is known as the *Hubble Parameter*, H. The relation between H and H_o is given in Appendix A. This factor sets the scale of the Universe as in measurements of large scale structure, there is an unknown factor h, where h is the *Hubble constant* in units of 100 km s^{-1} Mpc^{-1}.

Hubble's initial measurement of H_o yielded a high value of around 500 km s^{-1} Mpc^{-1}. However, errors were found in the calibration of the distance indicators. Most recent values, however, put H_o between 50 and 75 km s^{-1} Mpc^{-1}, e.g. Cepheid distances to the compact Leo-I group of galaxies implies H_o=67$\pm$7 km s^{-1} Mpc^{-1} (Tanvir, Ferguson &

*Through out this thesis we will use H_o=100 h km s^{-1}Mpc^{-1}

Shanks 1999) and recent results from the HST key project give H_o=68±7 km s^{-1} Mpc^{-1} (Gibson et al. 2000) using Cepheid and supernovae distances to seven galaxies.

The value of Hubble's constant has important implications for the age of the Universe. In a critical density Universe, defined in Appendix A, the age is found to be around 10 Gyr for a value of H_o = 70 km s^{-1} Mpc^{-1}. This is a problem as the ages of globular clusters have been estimated to be 16±2 Gyr (Renzini et al. 1996), older than the age of the Universe itself! One solution could be that Hubble's constant has been overestimated, perhaps the effects of dust or metallicity are more complicated than presently thought, leading to systematic underestimation of the distance measurements used to estimate H_o and hence an underestimation in the age of the Universe. Malmquist bias may also be affecting the measurements of H_o. Another possibility is that the mean density of the Universe is less than the critical density.

1.4 Aims of the Thesis

In order to answer some of the questions posed at the start of the Introduction, a better understanding of the structure and scale of the Universe is required. Our aim in this Thesis is to use redshift surveys of galaxies and QSOs to quantify the structure of the Universe and observations of galactic open clusters to check the calibration of the scale of the Universe. Below we outline the structure of the Thesis in more detail.

In Chapter 2, we review the status of redshift surveys of the local Universe and describe in detail the main surveys that are used in this Thesis. We also describe a series of simulations that are used throughout this Thesis.

In Chapter 3, we measure the power spectrum of galaxy clustering in the Durham/ UKST Galaxy Redshift survey. By using a combination of volume limited and flux limited samples, we aim to measure the clustering in the survey over as wide a range of scales, as accurately as possible. We compare the results with power spectra from other galaxy redshift surveys and quantify the differences between the various measurements. By applying a model for redshift space distortions to the APM real space power spectrum, we measure the one dimensional velocity dispersion and the parameter β. We also compare the power spectrum of the Durham/UKST survey to the power spectra of various models of structure formation in order to see which, if any, adequately match the clustering observed in the Universe.

In Chapter 4 we apply a different method of clustering analysis, counts-in-cells, to the

Durham/UKST and Stromlo-APM surveys. The aim is to measure directly the skewness of clustering on scales $\sim 20h^{-1}$Mpc. Again, we compare the results with models of structure formation with an aim to understanding the relationship between the distribution of the mass in the simulations and the galaxies in the real Universe.

In Chapters 5, 6 and 7 we present work in anticipation of the completion of the 2 degree Field (2dF) QSO survey. In Chapter 5, we draw mock 2dF QSO Survey catalogues from the ΛCDM *Hubble Volume* lightcone simulation, described in Chapter 2, that have approximately the same clustering as we expect the completed 2dF QSO Survey to have. We use these mock catalogues to determine how accurately and on what scales the correlation function and power spectrum of the 2dF QSO survey will be measurable. We also determine which methods and estimators will be optimal for measuring the clustering and which methods of error analysis give robust results.

In Chapter 6 we measure the power spectrum of QSO clustering from the current 2dF QSO survey. The QSO survey is only 2/5 complete and, due to the observing strategy adopted, there are large differences in the completeness of each area. However by taking these factors into account, an initial measurement is possible. The results are compared to galaxy and cluster power spectra and to different models of structure formation.

In Chapter 7, we construct models of the correlation function measured parallel and perpendicular to the line of sight, $\xi(\sigma,\pi)$. These models are then compared to $\xi(\sigma,\pi)$ measured from the mock catalogues to test if cosmological parameters can be constrained from the 2dF QSO survey. This method is known as the Λ test (Alcock & Paczyński 1979). We look to see if better constraints are possible if results from the evolution of the QSO-mass bias as a function of cosmology are combined with results from the Λ test.

In the earlier Chapters, we look at the clustering properties of galaxies and QSOs. As discussed earlier, the exact distance to a galaxy or QSO is not known as there is a factor h, where $h{=}H_0/100$ km s^{-1} Mpc^{-1}, in all the measurements. Many methods of determining Hubble's constant rely on the distance modulus to the Large Magellanic Cloud being accurately measured. Recently, Feast & Catchpole (1997) have challenged the previously accepted value of 18.50 as they obtained 18.70 from the Cepheid Period-Luminosity relation using trigonometrical parallaxes from Hipparcos to obtain the Cepheid distance and, hence, magnitude. This small change in the distance modulus to the LMC reduces the value of Hubble's Constant by 10%. Therefore, in Chapter 8, we use zero age main sequence fitting to stars in 11 Galactic Open Clusters containing Cepheids to check the calibration of the Cepheid Period-Luminosity relation, which is one method for calibrating

the distance modulus to the Large Magellanic Cloud.

Finally in Chapter 9, we draw conclusions and consider the future work required to further enhance our understanding of the structure and scale of the Universe.

Chapter 2

Galaxy and QSO Surveys

2.1 Types of Survey

During the past century, many different surveys of galaxies, and more recently of QSOs, have been carried out. These surveys can be characterised as angular or three dimensional. Angular catalogues show the distribution of galaxies as they appear projected onto the sky. These surveys typically contain many galaxies, for example the APM survey contains 2 million galaxies over an area of 4,300 square degrees (Maddox et al. 1990b). However, such surveys contain no direct information as to how far away each galaxy is; an apparently faint galaxy could be an intrinsically faint galaxy seen nearby or a bright galaxy at a large distance. Clustering statistics, such as the correlation function, obtained from these surveys are free from distortions caused by the peculiar motions of galaxies though.

In order to gain information about the three dimensional nature of the Universe, the distance to each galaxy is required. This is obtained by measuring the galaxy's redshift. Redshift is defined as the observed change in the frequency, or wavelength, of signals emitted from a source which is moving with respect to the observer. It is completely analogous to the Doppler effect, i.e.

$$1 + z = \frac{\lambda_o}{\lambda_e}, \tag{2.1}$$

where z is the redshift of the galaxy, λ_o is the wavelength of the observed light and λ_e is the wavelength of the emitted light. The redshift can then be converted into a distance as the recession velocity is proportional to distance (Hubble 1929), although this is only true locally. The full relation between redshift and distance is given in Appendix A.

As measuring redshifts is more time consuming than taking images of the sky, present redshift surveys contain far fewer galaxies than angular surveys. The largest completed galaxy redshift survey to date is the Las Campanas Redshift Survey (LCRS), containing around 26,000 galaxies (Shectman et al. 1996).

QSOs are believed to be the nuclei of active galaxies, fueled perhaps by an accreting black hole. The most important property of QSOs, for this work, is that they are intrinsically bright objects. QSOs have typical absolute b-band magnitudes of -23 which is roughly 3-4 magnitudes brighter than typical b-band galaxies. They can therefore probe the clustering at far higher redshifts than wide angle galaxy redshift surveys have so far been able to reach. For example, the median redshift of the Durham/UKST survey is 0.05 (Ratcliffe et al. 1998d) but the median redshift of the Durham/AAT QSO survey (Boyle 1986) is around 1.3, which, assuming a critical density Universe, translate to co-moving distances of $\sim 150h^{-1}$Mpc and $2040h^{-1}$Mpc respectively. However, some galaxy surveys, such as the CNOC Survey (Yee et al. 2000) have probed galaxy clustering out to $z = 0.6$ and surveys of Lyman break galaxies (Steidel et al. 1995) have probed galaxy clustering out to even higher redshifts of $z = 3$. These galaxy surveys currently only cover a small area of sky and, therefore, may not be representative of the full sample.

2.2 A Brief History of Surveys

2.2.1 Angular Surveys

Angular catalogues consist of the positions and magnitudes (in one or more wavebands) of galaxies as they appear projected on to the sky. Even though these surveys contain no information about how far away galaxies are, de Vaucouleurs (1948) was able to identify the existence of the Local supercluster from angular maps of the sky. The first angular catalogue that was statistically useful was the Lick survey (Shane & Wirtanen 1967). This is a large angular catalogue containing over one million images from which galaxies were identified by eye. This process was done over a long period of time by different observers and it has been suggested that there are observer dependent systematic effects in the final catalogue which can lead to spurious clustering results, particularly on large scales (de Lapparent, Kurtz & Geller 1986).

A major advance in angular catalogues came with the introduction of automatic plate measuring instruments such as the COSMOS machine (MacGillivray & Stobie 1984) and the Automatic Plate Measuring (APM) machine (Kibblewhite et al. 1984). These automatic measuring devices are able to measure the brightness, position and shape of each image. Using the COSMOS machine, the Edinburgh-Durham Southern Galaxy Catalogue (EDSGC Collins, Heydon-Dumbledon & MacGillivray 1988) was constructed with the aim of creating a large scale, homogeneous galaxy catalogue down to a magnitude

limit of $b_J \sim$20. The survey consists of 1.5 million galaxies over an area of 1500 square degrees. The APM catalogue (Maddox et al. 1990b) was constructed using the APM machine and detected 20 million images down to b_J=22 over an area of 4300 square degrees. From these images, a uniform sample of 2 million galaxies to b_J=20.5 was selected.

Angular catalogues have had an impact on our understanding of the large scale structure in the Universe, for example more large scale power was detected in the correlation function of the APM survey than expected from the standard Cold Dark Matter model (Efstathiou, Sutherland & Maddox 1990). The main use of these catalogues now is as input catalogues for redshift surveys. The EDSGC was used as the parent catalogue for the Durham/UKST survey and the APM survey was used for the Stromlo-APM survey and the 2dF Galaxy Redshift survey.

2.2.2 Galaxy Redshift Surveys

The first redshift survey consisted of only 25 galaxies and was carried out by Slipher in 1917. Arguably the most important redshift survey was the one carried out by Hubble in the 1920's which led to the discovery of the expansion of the Universe, as discussed in the Introduction. Catalogues of $\sim$ 1,000 galaxies with redshifts were published by Humason, Mayall & Sandage (1956) and de Vaucouleurs & de Vaucouleurs (1964) but it wasn't until the 1970's that galaxy redshift surveys were used to trace out the large scale distribution of galaxies. Gregory & Thompson (1978) confirmed the existence of superclusters from pencil beam surveys towards the Coma and Perseus clusters. Kirschner, Oemler & Schechter (1978) also carried out a series of pencil beam surveys, from which they measured the galaxy luminosity function. A deficiency of galaxies in part of the redshift distribution led to the detection of the Boötes void. The first wide angle survey was the revised Shapley-Ames Galaxy Catalogue by Sandage & Tammann (1981). This survey covers an area of around 20,000 square degrees and contains 1191 galaxies.

The first statistically useful redshift survey was probably the Centre for Astrophysics (CfA) survey (Huchra et al. 1983). This survey consists of 2,417 galaxies covering an area of 8,700 square degrees. When the redshifts of the galaxies are plotted against right ascension for a small slice in declination, voids and filaments are clearly visible. The most striking feature of the survey is the so-called 'stick man' which is the Coma cluster. The Durham/AAT survey (Peterson et al. 1986) consisted of only 329 field galaxies but observations were made to a fainter magnitude limit of $b_J \sim$17 to reduce any systematic effects due to local inhomogeneities in the clustering of galaxies. The largest completed

redshift survey to date is the LCRS (Shectman et al. 1996) which contains 26,000, R-band selected galaxies over six disjoint 1.5×80 degree slices in the north and south Galactic caps.

Surveys are also carried out in infrared wavebands. Infrared bright galaxies are predominantly found in the field rather than in clusters as the infrared emission is due to starlight being re-radiated by dust. In 1983, the Infrared Astronomical Satellite was launched and observed most of the sky at 12, 25, 60 and 100 μm. The catalogue created from these observations was used as the input catalogue for three other surveys considered in this Thesis. These are the QDOT Survey (Lawrence et al. 1999), the 1.2-Jy Survey (Fisher et al. 1995) and the PSCz Survey (Saunders et al. 2000). The amplitude of clustering from galaxies selected in the infrared tend to have a lower clustering amplitude than optically selected galaxies (Chapter 3).

All the completed galaxy redshift surveys considered further in this Thesis are summarized in Table 2.1. For a more comprehensive review of galaxy redshift surveys and clustering results from galaxy surveys see Strauss & Willick (1995) and Guzzo (1999).

There are currently two ambitious galaxy redshift surveys under way. These are the Two degree Field (2dF) Galaxy Redshift Survey and the Sloan Digital Sky Survey (SDSS). The 2dF survey (see, for example, Colless 1998) aims to measure the redshifts of 250,000 galaxies with a median redshift of 0.14 over a solid angle of ~1700 square degrees. The SDSS survey (see, for example, Gunn & Weinberg 1995) aims to measure the redshifts of nearly a million galaxies with a median redshift of 0.11 covering an area of ~10,000 square degrees. These surveys contain at least an order of magnitude more galaxies than existing galaxy redshift surveys and will provide unrivalled maps of the distribution of galaxies in the local Universe. More importantly, they will be sufficiently large that it will be possible to split the galaxy sample by colour, spectral type or luminosity to study the dependence of clustering on different galaxy properties and environments.

2.2.3 QSO Surveys

The first quasar was discovered by Schmidt (1963) and was detected due to its strong radio flux. However over 90% of QSOs do not have a detectable radio flux and are known as radio-quiet QSOs. These QSOs can be detected in three ways: by their extremely blue colour, ($U - B \lesssim -0.3$) (Sandage, Veron & Wyndham 1965, Schmidt & Green 1983, Shanks et al. 1983); by comparing recent photographic surveys with older surveys it is possibly to detect variations in the luminosities of QSOs (Koo, Kron & Cudworth 1986);

by selecting objects with broad emission lines from low dispersion spectra. This method was used to detect QSOs in the Large Bright Quasar Survey (LBQS, see below Hewitt, Foltz & Chaffee 1995).

Osmer (1981) and collaborators were the first to try and measure QSO clustering from surveys that contained over a hundred QSOs. No clustering was detected in the surveys on scales 100-3000h^{-1}Mpc and the number density in these surveys was such that the correlation function on smaller scales could not be measured. Shaver (1984) was able to demonstrate that QSOs were clustered on scales of $10h^{-1}$Mpc, using the inhomogeneous catalogue of QSOs collated by Veron-Cetty & Veron (1984). The clustering of the whole sample of QSOs was found to be similar to that of optically selected, present day galaxies. However, it was clear that a better defined sample of QSOs with a higher number density was required to measure QSO clustering more accurately (Osmer 1981).

The Durham/AAT QSO survey (Boyle 1986) consists of 392 QSOs to a magnitude limit of B=21, with redshifts in the range 0.3< z <2.2. This survey had a well defined selection criteria and small scale QSO clustering, consistent with that of optically selected galaxies, was again detected (Shanks et al. 1986, Shanks et al. 1987). The CFHT survey (Crampton, Schade & Cowley 1985) contains 215 QSOs, a large-scale QSO supercluster at $z = 1.1$ was detected in the survey but otherwise no significant clustering was measured from this survey. QSO surveys with brighter magnitude limits, such as the Homogeneous Bright QSO Survey (Cristiani et al. 1999) with QSOs observed in the range 15< B <18.75, found that the surface density of QSOs was higher that previously thought, this reduced the errors on the clustering statistics. Currently the largest completed QSOs survey is the LBQS (Hewitt, Foltz & Chaffee 1995). This survey consists of 1,053 QSOs covering an area of 454 square degrees. Croom & Shanks (1996) combined the Durham/AAT, CFHT and LBQS surveys and found significant clustering out to $30h^{-1}$Mpc scales.

All the above mentioned QSO surveys cover a far smaller area than galaxy redshift surveys but their large depth means they have a greater volume than the existing galaxy redshift surveys. There are currently around 10,000 published QSO redshifts from existing surveys (Veron-Cetty & Veron 1998) but these redshifts do not form a single, homogeneous sample. For a recent review of QSO clustering and surveys of QSOs see Cristiani (1999) and references therein.

As well as studying QSO clustering, there have been studies of the environments of QSOs. Yee & Green (1987) studied a sample of bright, radio-loud QSOs and found that at $z \sim 0.6$ they were found in environments as rich as Abell class 1 clusters. Ellingson,

Yee & Green (1991) observed a sample of 63 QSOs, of which around half were radio-loud and half were radio-quiet. They found that the excess population of galaxies around radio-quiet QSOs did not differ significantly from that expected around a normal galaxy. This suggests that radio-loud and radio-quiet QSOs are found in different environments. Smith, Boyle & Maddox (1995) also found that radio-quiet QSOs with redshifts in the range $0 < z < 0.7$ are found in similar environments as galaxies. Smith, Boyle & Maddox showed that the QSO-galaxy angular cross correlation function was indistinguishable from the APM galaxy-galaxy angular correlation function.

The 2dF and SDSS projects both have complementary QSO surveys. The 2dF QSO survey aims to measure the redshifts of 25,000 QSOs over an area of 740 square degrees. The SDSS survey aims to measure 100,000 QSOs, to a slightly brighter limit than the 2dF QSO Survey over a larger area. Both of these surveys will contain enough QSOs that the clustering amplitude of QSOs as a function of redshift will be measurable. Predictions for the 2dF QSO survey are made in Chapters 5, 6 and 7 and the 2dF QSO Survey is described in more detail in Section 2.3.4.

2.3 Details of Surveys used in this Thesis

Three surveys are used extensively in this Thesis. The Durham/UKST and the Stromlo-APM surveys are both galaxy redshift surveys. The magnitude limits of the two surveys are similar but the sampling rates adopted in each case from the parent angular catalogues are quite different. The other survey is the 2dF QSO survey, which has yet to be completed. Initial measurements of the correlation function are used in the construction of mock QSO catalogues in Chapter 5 and initial results for the power spectrum are given in Chapter 6.

2.3.1 The Durham/UKST Survey

The Durham/UKST survey consists of 2,501 galaxy redshifts, measured with the FLAIR fibre optic system (Parker & Watson 1995) on the UK Schmidt Telescope. The galaxies are sampled at a rate of 1 in 3 down to a magnitude limit of $b_J \sim 17$ from the parent EDSGC (Collins, Heydon-Dumbledon & MacGillivray 1988). The survey covers a solid angle of $\sim$ 1,450 square degrees and is more than 75% complete down to a magnitude limit of $b_J \sim 16.9$.

Full details of the construction of the Durham/UKST Survey, including the tests of

the accuracy of the measured redshifts and of the galaxy photometry can be found in Ratcliffe et al (1996, 1998d); see also Ratcliffe (1996).

2.3.2 The Stromlo-APM Survey

The Stromlo-APM survey is also optically selected, this time with a magnitude limit of $b_J = 17.15$. Observations were made on the 2.3m Mount Stromlo and Siding Spring Observatory telescope using a double beam spectrograph. The survey consists of 1,769 galaxies, sparsely sampled at a rate of 1 in 20 from the parent APM catalogue. The survey covers a solid angle of $\sim$ 4,300 square degrees. Full details are given in Loveday et al. (1996).

2.3.3 Comparison of the Durham/UKST and Stromlo-APM Surveys

In Chapter 3 and 4, we compare directly measurements of the clustering in the Durham/UKST and Stromlo-APM surveys. Here we compare the visual appearance and number density of these two surveys.

Figure 2.1, shows the positions of the galaxies in the Durham/UKST Survey (panels a-d) and Figure 2.2 shows the positions of the galaxies in the Stromlo-APM Survey that overlap with the Durham/UKST survey (panels e-h). Structures in the Durham/UKST Survey are clearly easier to pick out by eye, due to the six times higher sampling rate of this survey compared to that of the Stromlo-APM Survey. In the slices centered on $\delta = -30°, -35°$ and $-40°$, the Sculptor void is visible out to $60\,h^{-1}$ Mpc. The roof of this feature is seen in the $\delta = -25°$ slice.

The solid line in Figure 2.3 shows the radial number density of galaxies to a magnitude limit of $b_J \sim 17$, computed using the luminosity function parameters from the Durham/UKST survey (Ratcliffe et al. 1998a). The observed radial number density of galaxies, in bins of size $\Delta r = 10\,h^{-1}$ Mpc, is shown by the dashed line for the Durham/UKST Survey and by the dotted line for the Stromlo-APM Survey. The Durham/UKST number density (dashed line) lies a factor of three below the solid line due to the 1 in 3 sampling rate. The Stromlo-APM number density (dotted line) is approximately a factor of 20 below the solid line as galaxies were sampled at a rate of 1 in 20 but the magnitude limit of the Stromlo-APM survey is slightly fainter.

As well as comparing the results from the Durham/UKST survey to the results from the Stromlo-APM survey, we also compare against measurements from other surveys, taken from the literature. These surveys are summarised in Table 2.1. Some of these

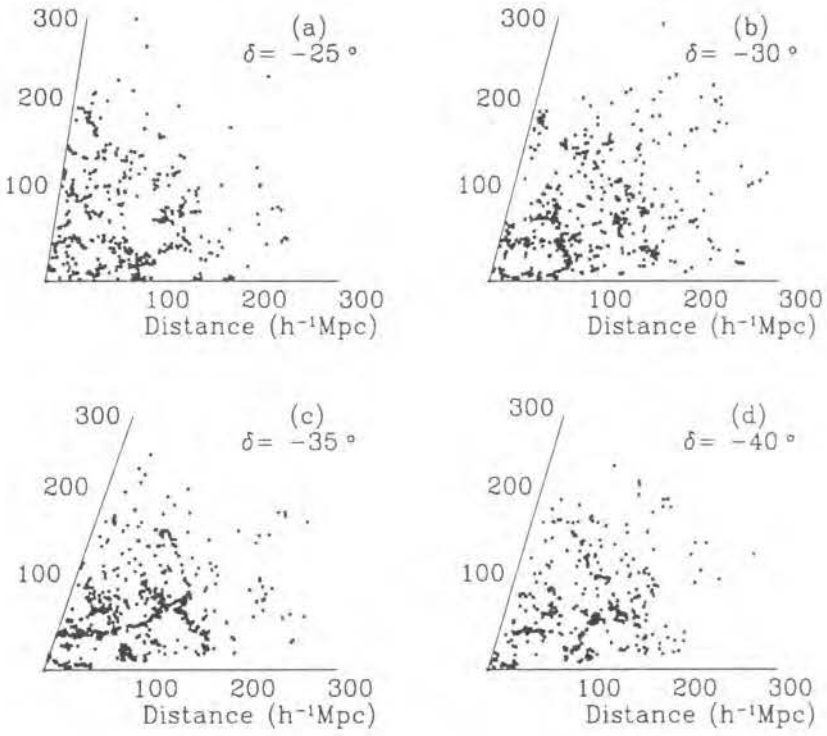

Figure 2.1: Galaxies in the Durham/UKST Survey are show in the plots (a-d). The declination slices are 5° thick and are centred on the declination shown in each panel. The RA ranges are: (a) $21^h27^m00^s < \alpha < 03^h29^m00^s$, (b) $21^h39^m30^s < \alpha < 03^h15^m30^s$, (c) $21^h48^m00^s < \alpha < 03^h24^m00^s$ and (d) $21^h53^m00^s < \alpha < 03^h41^m00^s$

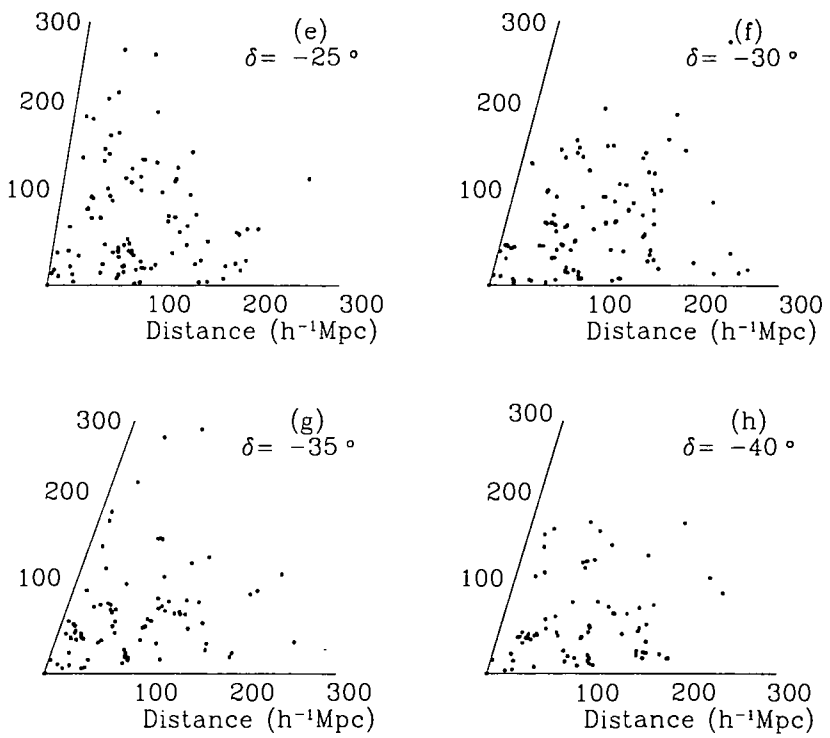

Figure 2.2: Galaxies in the Stromlo-APM Survey that overlap with the Durham/UKST survey are shown in the plots (e-h). The declination slices are 5° thick and are centred on the declination shown in each panel. The RA ranges of plots (e-h) are the same as those for plots (a-d) in Figure 2.1

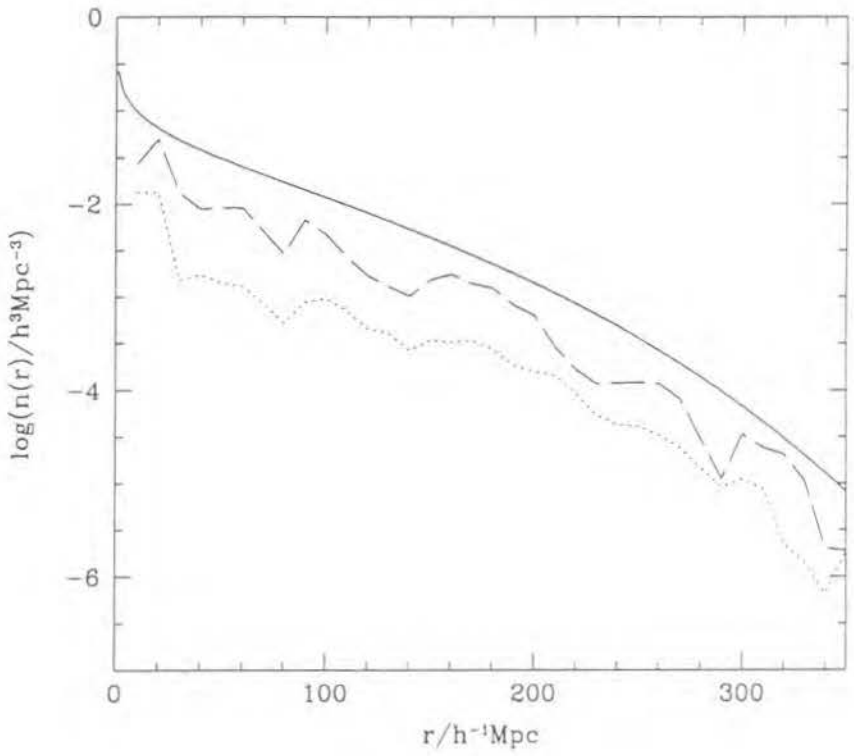

Figure 2.3: The solid line shows the radial number density of galaxies to a magnitude limit of $b_J \sim 17$, computed using the luminosity function of Ratcliffe et al. (1998a). The dashed line shows the observed number density of Durham/UKST galaxies, which are sampled at a rate of 1 in 3 from the EDSGC catalogue to this magnitude limit. The dotted line shows the radial number density of galaxies in the Stromlo-APM Survey, which are sampled at a rate of 1 in 20 from the parent APM catalogue to a magnitude limit of b_J=17.15.

Optical Surveys	N_{gal}	m_{lim}	Area (square degrees)	Reference
Durham/UKST	2501	$b_J \sim 17$	1450	Ratcliffe et al. (1998d)
Stromlo-APM	1769	$b_J = 17.15$	3700	Loveday et al. (1996)
Las Campanas	26418	R$\sim$17.7	700	Shectman et al. (1996)

IRAS Surveys	N_{gal}	Range (Jy)	Area (square degrees)	Reference
QDOT	2163	0.6-2	36000	Lawrence et al. (1999)
1.2-Jy	2663	1.2-1.936	36000	Fisher et al. (1995)
PSCz	15411	1.2	36000	Saunders et al. (2000)

Table 2.1: Main parameters for the Galaxy surveys used in this Thesis.

surveys are optically selected, whereas others are selected from the catalogue of sources compiled by the Infrared Astronomical Satellite (IRAS).

2.3.4 The 2dF QSO Survey

The 2dF QSO redshift survey aims to measure the redshifts of 25,000 QSOs over a redshift range of $0.3 \lesssim z \lesssim 3$. The spectra of the QSOs will be measured using the 2dF instrument on the Anglo-Australian Telescope (AAT). This instrument allows up to 400 spectra to be obtained simultaneously, allowing a large number of QSOs to be observed in a relatively short period of time. QSOs are selected in a homogeneous manner via U, B_J, R multi-colour selection and the survey will be more than 90% complete to $z \sim 2.4$. For more details see Croom et al. (1998a), Boyle et al. (2000) and Croom et al. (2000) and also `http://www.aao.gov.au/local/www/scroom`.

Two strips of the sky are being observed, each measuring $75° \times 5°$. One is centred at $\delta = -30°$, with $21^h40^m \lesssim \alpha \lesssim$ to 03^h15^m, close to the Southern Galactic Pole, while the other is centred at $\delta = 0°$ with $9^h50^m \lesssim \alpha \lesssim$ to 14^h50^m in the North Galactic Cap. The total area will be around 750 square degrees.

The aims of the survey include:

- measurement of the QSO correlation function and power spectrum over a wide range of scales;
- measurement of the evolution of QSO clustering over the redshift range $0.3 \lesssim z \lesssim 2.4$ to test models of structure formation;
- estimation of the cosmological constant via geometric distortions of the clustering pat-

tern.

These particular aims are discussed more in Chapters 5, 6 and 7.

Figure 2.4 shows the redshift, N(z), distribution of the 2,500 galaxies in the Durham/ UKST survey (filled histogram) as compared to the predicted distribution of the 25,000 QSOs in the 2dF QSO survey (open histogram) scaled up from the N(z) distribution of the first ~4,000 QSOs to be observed. This Figure clearly demonstrates the vast range of redshifts that QSO surveys are able to probe as compared to wide angle galaxy surveys. As of January 2000, $\sim$ 9,000 QSOs have been observed in the 2dF QSO survey. The redshift distribution of these QSOs is shown in Figure 2.5. The comoving distances are calculated assuming a cosmology with $\Omega_{\rm m}$=1 (see Appendix A for a definition of $\Omega_{\rm m}$, Ω_Λ and the relationship between cosmological parameters and comoving distance). The two strips at the top and bottom of the plot show the distribution of QSOs as they would appear on the sky. The plot is courtesy of the 2dF QSO team.

2.4 Simulations and Models of Structure Formation

Simulations of the Universe are used in this Thesis to directly test models of structure formation against observations of the real Universe, as a means to obtain errors on clustering statistics and to develop and test clustering analysis techniques in a controlled situation.

Directly comparing clustering obtained from models of structure formation with observations of clustering in the Universe is probably the most important use of simulations as this potentially allows cosmological parameters to be constrained. The errors on currently available measurements of clustering allow some level of discrimination between various models (see Chapters 3 and 6 for an example). However, surveys in progress, such as the 2dF and SDSS, will be able to place far tighter constraints on the models.

Simulations also allow us to gain insight into the statistical significance of clustering observed in the Universe. A large number of independent mock catalogues, which contain the same number of objects and have the same geometry and radial selection function as a survey, can be used to derive the errors on measured clustering statistics. This approach is valid as long as the clustering pattern of the simulation is similar to that of the observed universe. If not, a biasing prescription can be applied to the simulation to force the clustering to match the observed clustering pattern. This is one approach that we use to estimate the errors on the power spectrum (Chapters 3 and 6) and on the two-point correlation function (Chapter 5).

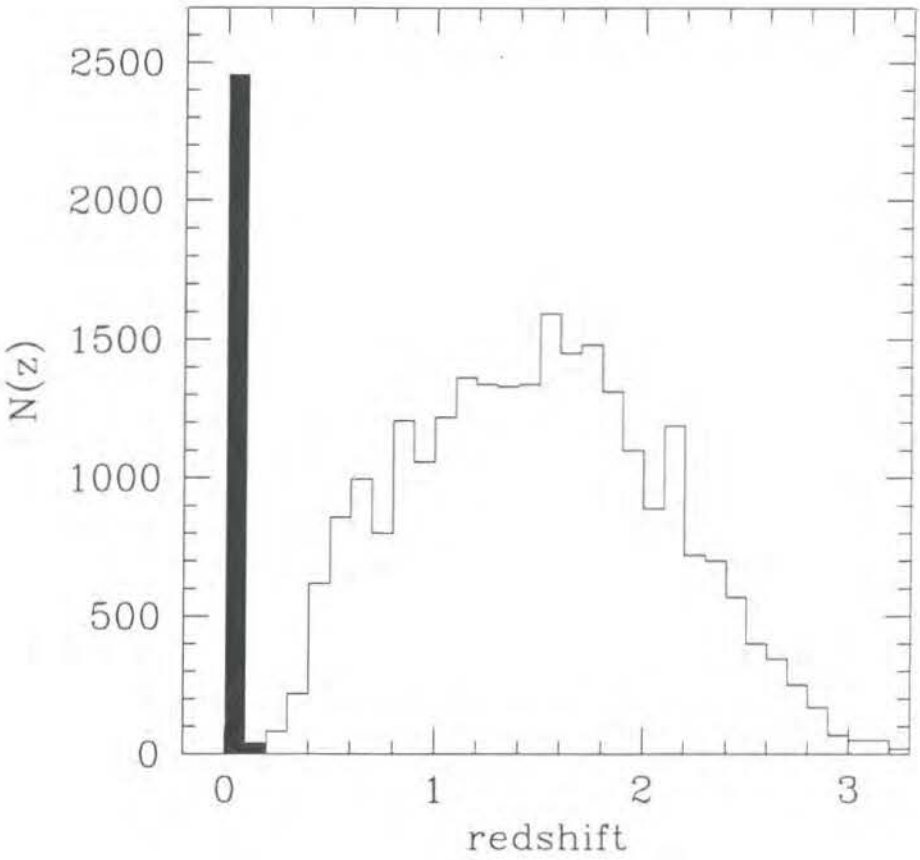

Figure 2.4: The open histogram shows the distribution of QSOs versus redshift measured from the first 4,000 QSOs to be observed and scaled so that the histogram contains 25,000 QSOs. The filled histogram shows the distribution of galaxies in the Durham/UKST survey. The range of redshifts probed by QSO surveys is far larger than that of wide angle galaxy surveys.

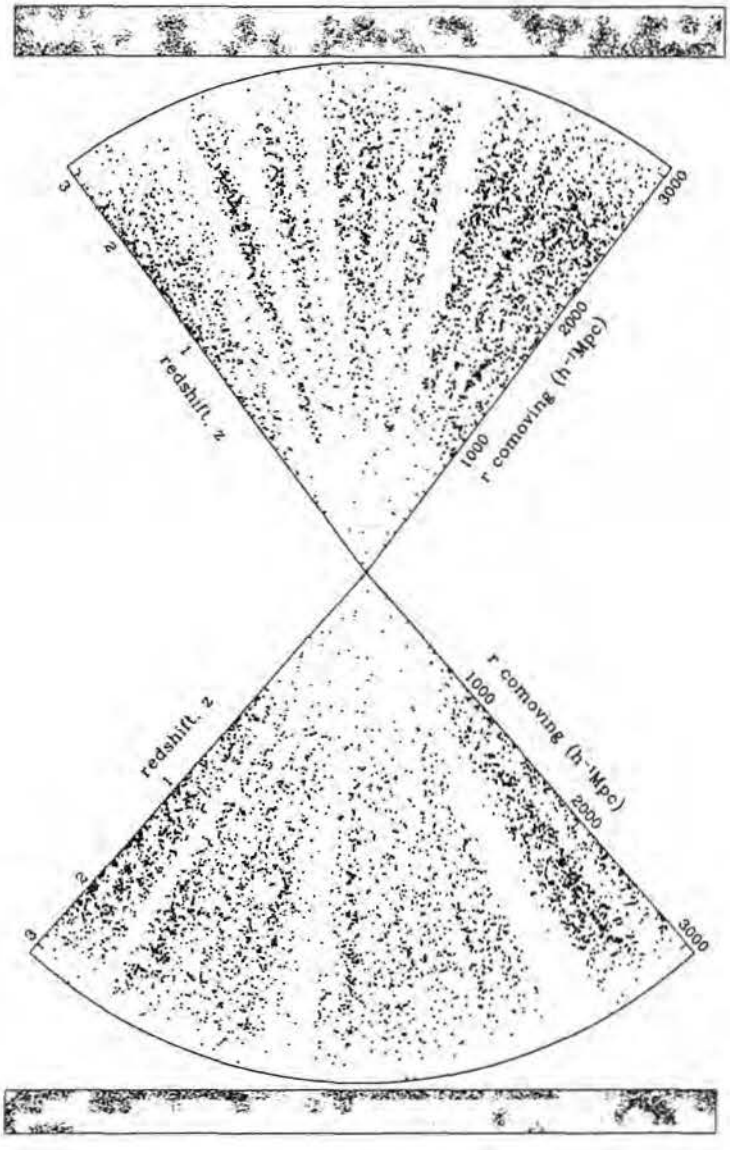

Figure 2.5: The status of the 2dF QSO survey as of January 2000. $\sim$ 9,000 QSOs have so far been observed. This plot is courtesy of the 2dF Team. The comoving distances are calculated from the redshifts assuming an Ω_m=1 cosmology. The two strips show the distribution of QSOs as they would appear on the sky.

One set of simulations, the *Hubble Volume* simulations, is used extensively in this Thesis and is described in section 2.4.2.

2.4.1 N-body Simulations and the Cold Dark Matter Scenario

The first computer N-body simulations were run by Peebles in 1969, but it wasn't until the 1980's that simulations began to make a significant impact upon studies of large scale structure. Until then, there was a lack of physically motivated initial conditions to use in the simulations and a lack of efficient techniques for handling a large number of particles that develop into a highly clustered distribution.

The size of N-body simulations has grown rapidly over the last few decades, due mainly to the increased power of computers, but also due to improvements in programming techniques. Figure 2.6 shows how the size of cosmological N-body simulations (defined by the number of particles in the simulation rather than the volume of the simulation) has grown over the last 30 years as compared to the growth in the size of optical redshift surveys over the same period. Over the last 10 years, the simulations have grown at around the same rate as the redshift surveys but the simulations contain a factor of 10,000 more particles; this resolution is necessary if we wish to follow the formation of individual dark matter halos in the simulations.

Simulations of large scale structure are performed within the context of gravitational instability. Other ingredients that must be specified are

- the material content of the Universe;
- the amplitude and pattern of initial fluctuations;
- the cosmology.

The material content of the Universe must be specified in order to determine the shape of the power spectrum at recombination. (The power spectrum is defined in Appendix B.) Baryonic material must make up at least part of the matter content of the Universe, though the precise amount is unknown. Constraints from nucleosynthesis suggest that $0.01 \lesssim \Omega_{\rm b} h^2 \lesssim 0.024$ (Tytler, Fan & Burles 1996), implying, for example, that $0.04 \lesssim \Omega_{\rm b} \lesssim 0.1$ if H_0=50km s^{-1} Mpc^{-1}. Many measurements of $\Omega_{\rm m}$ give values larger than 0.1 (see, for example, Eke, Cole & Frenk 1996, Perlmutter et al. 1999, Lange et al. 2000) so what does the rest of the mass consist of? One suggestion is that neutrinos have a mass of around 30eV. This is known as the hot dark matter (HDM) scenario as the neutrinos move at relativistic velocities at the epoch of matter-radiation equality. Another suggestion is that unknown, non-baryonic particles, which move at non-relativistic velocities and which

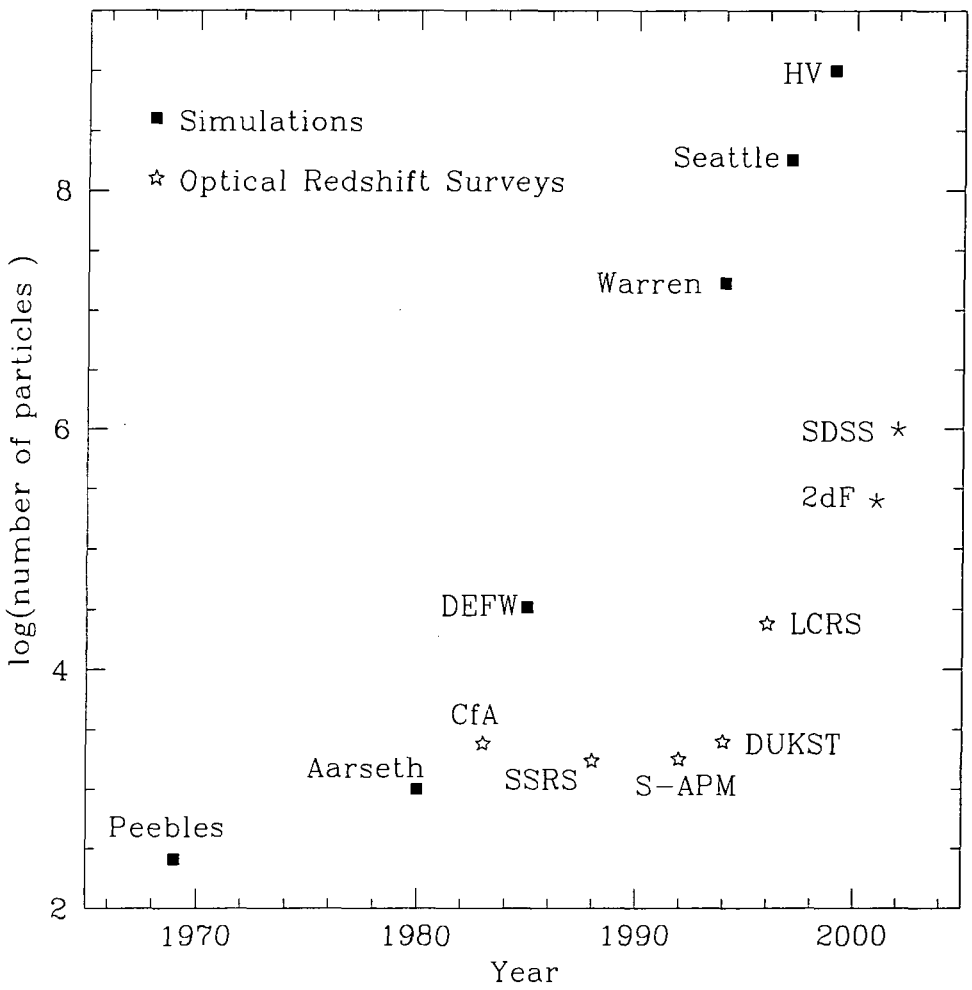

Figure 2.6: The increase in the number of particles in simulations (filled squares) and the increase in the number of measured redshifts in galaxy surveys (open stars) as a function of time. The values for the 2dF and SDSS Surveys are shown by a different star symbol as we show the expected numbers as the surveys have not been completed.

probably interact only through gravity, make up the missing mass, this is known as the cold dark matter (CDM) scenario. Simulations of both scenarios have been carried out and the clustering in CDM simulations matches the observations better than the clustering in HDM simulations (White, Frenk & Davis 1983, Davis et al. 1985). The CDM model is therefore currently the most popular choice for the material content of simulations. There are still some problems with the CDM model though, for example the velocities of particles in the simulations are too high compared to those observed for galaxies (but see Benson et al. 2000a).

The normalisation of the fluctuations in the simulations is defined by a parameter known as σ_8, where σ_8^2 is defined as the normalised linear theory variance in the mass contained within spheres of radius $8\,h^{-1}$ Mpc at $z = 0$; $\sigma_8^2 =< M - \bar{M} >^2 / \bar{M}^2$. The value of σ_8 for the mass was unknown until the early 1990's but it can now be determined in two ways. One method is from fluctuations in the temperature of the CMBR (Smoot et al. 1992). The shape and amplitude of the angular power spectrum of temperature fluctuations provide information on both the spectral index and amplitude of the primordial power spectrum, given an assumption about the contribution of density fluctuations to the temperature power spectrum. This normalises the power spectrum on scales of $\sim 1000h^{-1}$Mpc. If the shape of the power spectrum is also specified, then the amplitude of the power spectrum at $8\,h^{-1}$ Mpc can be determined, yielding σ_8. The second method uses the abundance of hot X-ray clusters (White, Efstathiou & Frenk 1993, Eke, Cole & Frenk 1996). The mass of X-ray clusters is related to the temperature and, through the theory of Press & Schechter (1974), the abundance of X-ray halos can be calculated. This then allows the variance in the mass contained within spheres of radius $8\,h^{-1}$ Mpc to be defined, assuming that the initial density fluctuations are Gaussian.

If the material content of the Universe is specified as CDM, which is the case for all the simulations considered in this Thesis, then the shape of the power spectrum can be parameterised by a factor Γ, where $\Gamma = \Omega_{\rm m} h$ in purely CDM models. If the simulation contains baryons, then $\Gamma = \Omega_{\rm m} h$ is only an approximation as the power spectrum is damped on small scales. A more accurate formula for Γ is given by Sugiyama (1995). If models are cluster normalised, (i.e. normalised on small scales) then models with smaller values of Γ have more large scale power than models with larger values of Γ, see Figure 2.7(a). A further parameter is the index of the primordial power spectrum, n. In all the CDM models used here, n is assumed to be 1, which is predicted by inflation and is consistent with results from CMBR experiments (Lange et al. 2000). On large

scales, this is the slope of the power spectrum. However, the power spectrum from CDM simulations on smaller scale has a different slope, $n \sim -2$ to -3. The power spectrum therefore turns over from $n = 1$ to $n = -2$ to -3. The exact scale on which the power spectrum slope changes from positive to negative is referred to as the turnover. The shape of the power spectrum changes as the size of the cosmological horizon changes with time. At early times, the horizon is small. The growth of fluctuations on scales smaller than the horizon is suppressed because of the Mésáros effect (Mésáros 1974), as the pressure of the gravitationally dominant photon fluid causes acoustic oscillations in density perturbations, which prevents fluctuations from growing until dark matter becomes the dominant gravitational component. As the horizon grows, larger scale fluctuations enter the horizon, causing damping of the power spectrum on ever larger scales. This continues until the epoch of matter-radiation equality, after which the Mésáros effect is no longer effective. The shape of the power spectrum on small scales thus depends on the relative length of time which fluctuations of different scales spent within the horizon before the epoch of matter-radiation equality. After matter-radiation equality, the Mésáros effect no long damps the fluctuations and fluctuations on the largest scale can continue to grow, producing a power spectrum with slope $n = 1$ on the largest scales.

These parameters are summarized in Figure 2.7. The solid lines show different values of Γ, with power spectra with Γ=0.1 and 0.5 labeled. The other power spectra vary between 0.1 and 0.5 in steps of 0.1. The upper panel shows power spectra that are cluster normalised. The lower panel shows power spectra that are COBE normalised. The initial slope of the power spectrum is $n = 1$ in all cases, shown by the dashed line. On smaller scales, n decreases. The scale on which the turnover occurs depends on Γ. Only models with $\Gamma \sim$0.2 can match both the COBE and cluster normalisation.

The values of the cosmological parameters, which determine the expansion rate and the growth of fluctuations, are as yet unknown. Normally a range of CDM models are considered. The original CDM model has parameters Ω_m=1, H_o=50 km s^{-1} Mpc^{-1}, a (cluster) normalisation of σ_8=0.52, and a shape parameter of $\Gamma = \Omega_m h = 0.5$ and is known as Standard CDM (SCDM). This model contains too little structure on large scales (Efstathiou, Sutherland & Maddox 1990). One solution is to allow $\Omega_m < 1$ and have either a spatially flat Universe with a cosmological constant (ΛCDM) or an open Universe without a cosmological constant, (OCDM). In both cases $\Gamma \sim$0.2 as $\Omega_m \sim 0.3$ and H_o=70 km s^{-1} Mpc^{-1} are the approximate parameters. The ΛCDM model is currently the most favoured model; a flat cosmology agrees with current observations from CMB

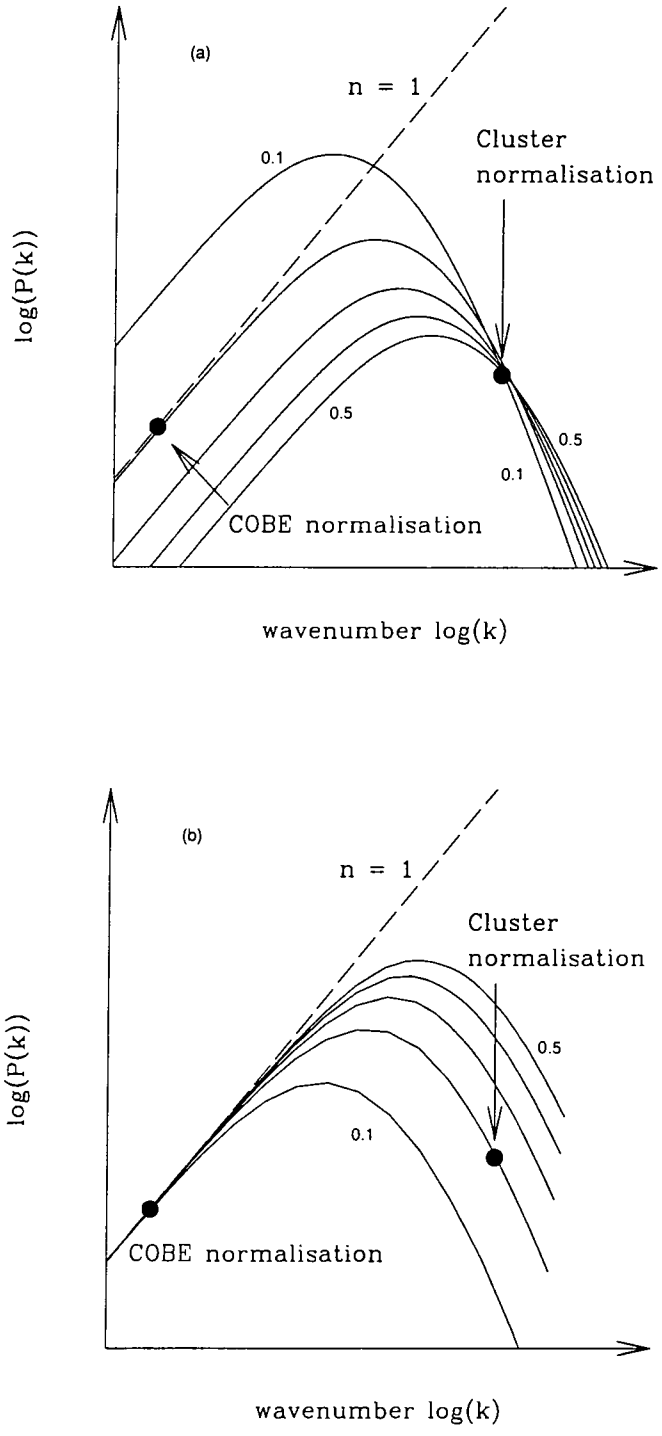

Figure 2.7: Shown above are pure CDM power spectra with different values of Γ, the power spectra with Γ=0.1 and 0.5 are labeled and the other power spectra have values of Γ that vary between 0.1 and 0.5 in steps of 0.1. Panel (a) shows power spectra that are cluster normalised and panel (b) shows power spectra that are COBE normalised. When $\Gamma \sim 0.2$, the power spectrum can match both constraints. The dashed line shows a power law with $P(k) \propto k$, which is the initial power spectrum shape in all the models, consistent with results from CMBR experiments (Lange et al. 2000).

experiments (de Bernardis et al. 2000, Balbi et al. 2000) and matches both the COBE and cluster normalisations, as discussed above. However, Ω_m=1 remains a popular choice for the cosmology as there is no convincing physical explanation of the cosmological constant. It is still possible for Ω_m to equal unity if the shape of the power spectrum is altered to match both the cluster and COBE normalisation, i.e. Γ=0.2 in this model. This change in the shape of the power spectrum could be achieved by postulating a massive particle, whose decay produces an additional contribution to the radiation density of the universe. This would delay the epoch of matter radiation equality (White, Gelmini & Silk 1995). This model is known as τCDM.

2.4.2 The Hubble Volume Simulations

The *Hubble Volume* Simulations * are a series of simulations run by the Virgo Consortium and to date are the largest N-body simulations ever. Each simulation consists of one billion particles. Here we describe the details most pertinent to the analysis contained in this thesis. For more details of the simulations run by the Virgo Consortium, see `http://star-www.dur.ac.uk:80/ ~frazerp/virgo/virgo.html`, Jenkins et al. (1998) and Evrard et al. (2000)

The Durham/UKST and Stromlo-APM surveys are essentially carried out at redshift zero. The median galaxy in each survey is at $z \sim 0.05$. If the comoving distance to this galaxy is calculated assuming a critical density, $\Omega_m = 1$ cosmology or a cosmology where $\Omega_m = 0$, Ω_Λ=1, there is only a difference of around 3% between the two obtained values. The low median redshift means the time slice outputs at $z = 0$ are suitable simulations for mock Durham/UKST or Stromlo-APM catalogues.

There are two different $z = 0$ simulations. One has a τCDM cosmology with $\Omega_m = 1$, $H_0 = 50$km s^{-1} Mpc^{-1} but with a spectral shape parameter of Γ=0.21 rather than the standard value of Γ=0.5 and a normalisation of σ_8=0.6. The volume of the $z = 0$ τCDM simulation is $2000\times2000\times2000h^{-3}$Mpc^3, allowing at least 10,000 independent Durham/UKST surveys to be extracted from the simulation. Typically 40 mock catalogues are sufficient to estimate the errors on the power spectrum of the Durham/UKST

*The Hubble Volume simulations were performed by the "Virgo consortium for cosmological simulations". This is an international collaboration involving universities in the UK, Germany and Canada. The members of this consortium are: J. Colberg, H. Couchman, G. Efstathiou, C. Frenk (PI), A. Jenkins, A. Nelson, J. Peacock, F. Pearce, P. Thomas, and S. White. G. Evrard is an associate member. The Hubble Volume simulation was carried out on the Cray-T3E at the Max-Planck Rechen Zentrum in Garching.

survey, see Chapter 3. The other simulation has a ΛCDM cosmology with Ω_m=0.3 and Ω_Λ=0.7, normalised to σ_8=0.9 and H_0 = 70km s^{-1} Mpc^{-1}. The volume of the $z = 0$ ΛCDM simulation is $3000\times3000\times3000h^{-3}Mpc^3$.

The 2dF QSO survey extends out to a redshift of $z \sim 3$. In order to make realistic mock catalogues, a simulation where the dark matter particles are output along an observer's past lightcone is required. The ΛCDM simulation, with parameters as above, has one deep lightcone output which covers an area of 15 × 75 degrees and extends out to $z \sim 4$. We use this simulation to create mock 2dF QSO catalogues to test estimators of clustering statistics in advance of the completion of the 2dF QSO Survey.

Chapter 3

Power Spectrum Analysis of the Durham/UKST Galaxy Redshift Survey

3.1 Introduction

Measuring the primordial power spectrum of density fluctuations in the Universe is of fundamental importance in the development of a model of the formation of large scale structure. The shape and amplitude of the power spectrum contains information about the nature of dark matter and the relative densities of dark matter and baryons. Several obstacles prevent a direct measurement of the primordial power spectrum from surveys of the local Universe. Structures are mapped out by galaxies and these may be biased tracers of the underlying mass distribution (Davis et al. 1985). Furthermore, the relation between fluctuations in the galaxy and mass distributions could be a function of scale and this needs to be addressed with a model for galaxy formation (e.g. Benson et al. 2000b). The pattern of clustering is also distorted when galaxy positions are inferred from their redshifts. This is due to a contribution to the observed redshift from the peculiar motion of the galaxy, that arises from inhomogeneities in the local gravitational field, in addition to the contribution from the Hubble flow (Kaiser 1987, Peacock & Dodds 1994).

Measurements of galaxy clustering have improved dramatically in the last ten years with the completion of several large galaxy surveys. The infrared selected QDOT redshift survey (Efstathiou et al. 1990) and the optical, angular APM Survey (Maddox et al. 1990b) were the first to demonstrate that there is more power in the galaxy distribution on large scales than expected from the standard cold dark matter (CDM) theory of structure formation. This led to variants of the standard CDM picture being considered.

The power spectrum has become the favoured statistic for quantifying galaxy cluster-

ing. This is despite the development of improved estimators for the two-point correlation function (Hamilton 1993, Landy & Szalay 1993), described in more detail in Chapter 5. Both statistics are affected by uncertainties in the mean density of galaxies, however these uncertainties affect the correlation function on all scales whereas they only affect the power spectrum on large scales (Cole, Fisher & Weinberg 1995). The power spectrum is also the quantity predicted directly by theory. Errors in the power spectrum are essentially uncorrelated before the mixing of different Fourier modes due to the convolution of the power spectrum of the galaxy clustering with that of the survey window function. Power spectra are also usually estimated using a Fast Fourier Transform (FFT) and are therefore relatively quick to compute. Recent theoretical work (Tegmark et al. 1998) has demonstrated that power spectrum analysis can be extended to adjust for various systematic effects and biases in the data, such as obscuration by dust or the integral constraint, which is discussed in Section 3.3. However, in general these corrections require an assumption about the form of the underlying power spectrum and are therefore model dependent. For this reason, and because the more advanced analysis outlined by Tegmark et al. (1998) has yet to be applied to a large, optically selected galaxy survey to enable a comparison, the approach developed by Feldman, Kaiser & Peacock (1994) and Tadros & Efstathiou (1996) is adopted.

We apply power spectrum analysis to the Durham/UKST galaxy redshift survey, described in Chapter 2. The clustering of galaxies in this survey has been studied using the two-point correlation function by Ratcliffe et al. (1996) and Ratcliffe et al. (1998b) and the magnitude of redshift space distortions were considered in Ratcliffe et al. (1998c). Although the two point correlation function is the Fourier transform of the power spectrum, the same is not true of a noisy estimate of the two-point function. In addition to studying a flux limited sample, in which the galaxies are weighted such that the variance in the power spectrum estimate is minimised, volume limited samples are also considered, in which all galaxies are given equal weight.

The construction of different subsamples of the survey for power spectrum analysis is considered in Section 3.2. Power spectrum estimators are tested using mock catalogues drawn from a large numerical simulation of clustering in Section 3.3. In Section 3.4 the results are presented and these are compared with other Surveys in Section 3.5. The implications for models of large scale structure formation are discussed in Section 3.6 and conclusions are given in Section 3.7.

3.2 Power Spectrum Analysis

3.2.1 Sample definition

The observational properties of the Durham/UKST survey are discussed in Chapter 2. Two types of galaxy sample are used in the power spectrum analysis: (i) flux-limited and (ii) volume limited. In order to estimate the power spectrum of galaxy clustering in these samples, we also need to construct sets of unclustered points with the same radial and angular selection; this process is described in Section 3.2.2.

Flux-limited sample

In this case, all galaxies with measured redshifts are used. A weight is assigned to each galaxy to take into account the radial selection function of the survey, shown for a magnitude limit of b_J=17 in Chapter 2, Figure 2.3. We adopt the form of the weight proposed by Feldman, Kaiser & Peacock (1994), which minimises the variance in the estimate of the power spectrum:

$$w(r_i) = \frac{1}{1 + n(r_i)P(k)} . \tag{3.1}$$

Here $n(r_i)$ is the mean galaxy density at the position of the $i^{\rm th}$ galaxy. This is calculated by integrating over the luminosity function of the survey, taking into account the sampling rate. There is a slight difference in the magnitude limit of each Schmidt plate in the Durham/UKST Survey (see Figure 1 of Ratcliffe et al. 1998b), so a separate radial weight function is computed for each plate. Ideally, one should use the true power spectrum in the weight given by equation 3.1. However, the results are fairly insensitive to the exact choice of power spectrum. Following the approach taken by Feldman, Kaiser & Peacock (1994) and by Tadros & Efstathiou (1996), we adopt a range of constant values of $P(k)$ that are representative of the amplitude of the power spectrum over the wavenumbers of interest. We define the depth of the sample as the distance for which the radial weight function $w(r) = 0.5$. For our choices of constant power in equation 3.1, this gives depths in the range 200–320 h^{-1} Mpc. The power spectrum analysis of the flux limited catalogue therefore probes volumes in the range 1.2–4.9 $\times$ $10^6 h^{-3}$ Mpc^3.

Volume-limited samples

The galaxies in a volume limited sample are brighter than the apparent magnitude limit of the survey when placed at any redshift up to that used to set the volume limit, $z \leq z_{\rm max}$.

Hence, as well as requiring that a galaxy has a redshift $z \leq z_{\rm max}$, the absolute magnitude of the galaxy must be brighter than:

$$M_{\rm crit} = m_{\rm lim} - 25 - 5\log_{10}\left[d_l(z_{\rm max})/h^{-1}{\rm Mpc}\right] - k(z_{\rm max}) \qquad (3.2)$$

where $m_{\rm lim}$ is the magnitude limit of the survey and we use the k-correction, $k(z)$, given by Ratcliffe et al. (1998a). Again, the different plate magnitude limits are taken into account, so for a given redshift limit, $z_{\rm max}$, the critical absolute magnitude varies slightly from plate to plate. We compute the luminosity distance, d_l, assuming an $\Omega_{\rm m} = 1$ cosmology, although our results are insensitive to this choice due to the relatively low redshifts of Durham/UKST galaxies.

In the Durham/UKST Survey, the number of galaxies in a volume limited subset peaks at a redshift of $z_{\rm max} = 0.06$ (Figure 3.1). There are 522 galaxies in this sample. There is a slightly smaller peak for a sample limited at $z_{\rm max} = 0.04$. This feature is particularly strong on the plate centered on δ=-35°, see Figure 2.1 in Chapter 2. The same peaks are also seen in volume limited subsamples of the Stromlo-APM Survey when attention is restricted to those galaxies that overlap with the Durham/UKST Survey. The dotted lines in Figure 3.1 are theoretical curves calculated by integrating over the luminosity function. The volume limited samples considered have maximum depths in the range 120–230 h^{-1} Mpc, and thus sample volumes of 0.2–1.8 $\times$ $10^6\,h^{-3}\,{\rm Mpc}^3$.

3.2.2 Survey geometry and radial selection function

The power spectrum measured directly from a galaxy survey is a convolution of the true power spectrum of the galaxy clustering with that of the survey window function. The power spectrum of the survey window function is estimated by placing a large number of unclustered points, typically on the order of 100,000, within the angular area covered by the survey. The random points must also have the same radial selection function as the galaxy sample under consideration.

To construct the random catalogue, points are distributed randomly within a box that is larger than the survey with the observer placed at one corner of the box. Particles with angular coordinates that do not lie within the survey are immediately rejected, as are points that lie at depths greater than the volume limit of the sample under consideration. This method generates a random catalogue that has a radial selection function that varies little as a function of distance from the observer out to the maximum depth of the sample under consideration, as found for volume limited samples of galaxies.

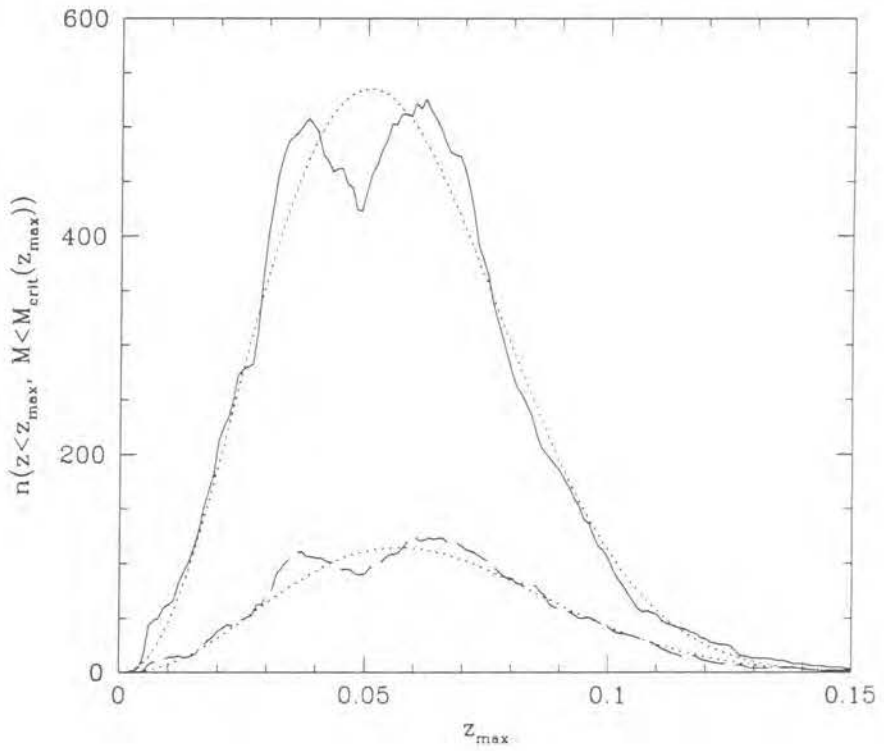

Figure 3.1: The number of galaxies in volume limited samples as a function of the redshift used to define the volume limit, z_{max}. The solid line shows the number of galaxies in volume limited samples drawn from the Durham/UKST Survey. The dashed line shows the number of galaxies from the Stromlo-APM Survey that satisfy the volume limit constraints, and which lie on the same Schmidt plates. The number of galaxies in both cases peaks for a sample limited at $z_{max} = 0.06$ – there is also a strong feature that can be seen in the two catalogues around $z_{max} = 0.04$. The dotted lines show the expected number of galaxies obtained by integrating over the luminosity function, taking into account the different sampling rates of the two surveys.

The radial selection function for the flux limited sample varies as a function of scale, see Figure 2.3 in Chapter 2. As before, we randomly generate a particle position in a box that is larger than the survey. We again reject particles that lie outside the survey geometry. We match the radial selection function of the flux limited sample by assigning an absolute magnitude to each particle selected randomly from the luminosity function. From the particle's position, we calculate the apparent magnitude and this is checked to ensure it is brighter than the plate limit. If not, the particle is rejected. The remaining particles then have the same angular and radial selection function as the flux limited sample of galaxies. We ensure that we have at least ~100,000 random points within the survey window of the sample under consideration.

Figure 3.2 shows the power spectrum of the Durham/UKST Survey window function for various volume limited and flux limited samples. The top panel shows the power spectra of the window function for different volume limited samples. The width of the window function power spectrum decreases as the volume limit adopted increases. Figure 3.2(b) shows the window function power spectra of flux limited samples. As the value of the power used in equation 3.1 is increased, the flux limited sample has a larger effective depth and so the width of the window function is reduced. There is a relatively small change in the width of the survey window function when different samples of the data are considered. Defining the effective width of the window function as the wavenumber at which the power spectrum of the window function falls to half its maximum value, we obtain $\delta k \sim 0.015\, h\,\mathrm{Mpc}^{-1}$. At wavenumbers smaller than this, the estimates of the power will be strongly correlated. For both flux limited and volume limited samples, the window function power spectrum is a very steep power law at wavenumbers $\log(k/h\mathrm{Mpc}^{-1}) \geq -1.20$, varying as k^{-4}.

3.2.3 Power spectrum estimation

The power spectrum estimator that we employ is a generalisation of that given by equation 12 of Tadros & Efstathiou (1996), see also Sutherland et al. (1999), to include the analysis of flux limited samples. We do not reproduce all the details of their derivation here.

The Fourier transform of the observed galaxy density field, within a periodic volume V, is given by

$$\hat{n}_o(\mathbf{k}) = \frac{1}{V}\sum_i w_{\mathrm{gal}}(\mathbf{x}_i) e^{i\mathbf{k}.\mathbf{x}_i}\,, \tag{3.3}$$

where the weight function $w_{\mathrm{gal}}(\mathbf{x}_i)$ depends upon the type of galaxy sample under consid-

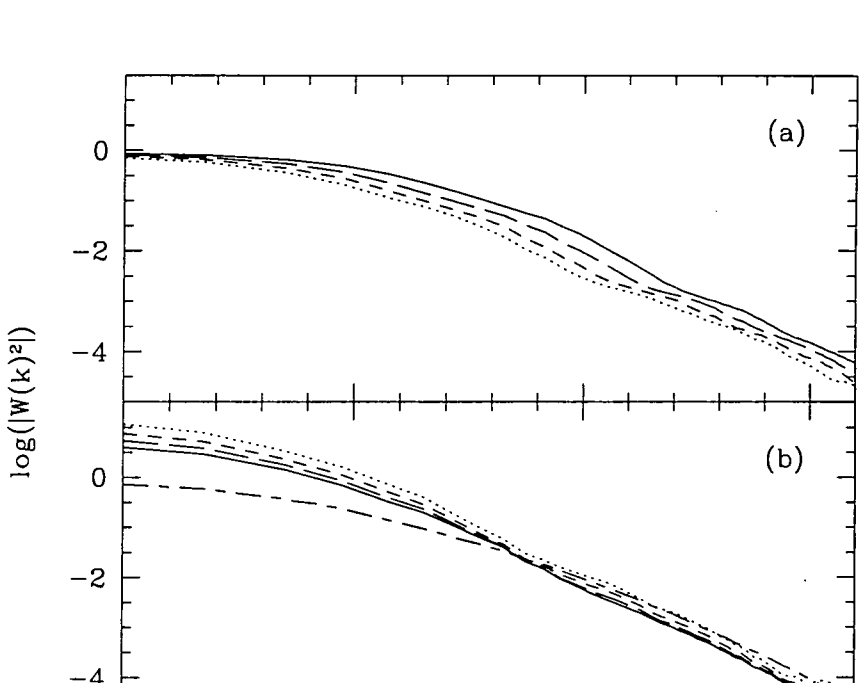

Figure 3.2: The two panels show the power spectrum of the window function for different samples extracted from the Durham/UKST survey. In (a), the samples are volume limited with a maximum redshift of $z_{\max}$=0.05, 0.06, 0.07, 0.08 reading from top to bottom. In (b), we plot the power spectrum of the survey window function for flux limited samples. The weights applied are computed assuming P=32000, 16000, 8000, 4000 and $0\,h^{-3}\,\mathrm{Mpc}^3$ reading from top to bottom at $\log k = -1.5$. For wavenumbers $k \geq 0.06\,h\,\mathrm{Mpc}^{-1}$, the window function power spectrum is a steep power law, $\propto k^{-4}$.

eration. For the case of a volume limited sample, $w_{\rm gal}(\mathbf{x}_i) = 1$ for a galaxy which satisfies the criteria given in Section 3.2.1 and $w_{\rm gal}(\mathbf{x}_i) = 0$ otherwise. For a flux limited sample, $w_{\rm gal}(\mathbf{x}_i)$ is given by equation 3.1.

The Fourier transform of the survey window function is approximated by:

$$\hat{W}_e(\mathbf{k}) = \frac{1}{V}\sum_i w_{\rm ran}(\mathbf{x}_i)e^{i\mathbf{k}.\mathbf{x}_i}, \tag{3.4}$$

where $w_{\rm ran}$ is the weight assigned to one of the unclustered points used to trace out the survey volume (note that the definition we have adopted for the Fourier transform of the survey window function differs by a factor of $1/\bar{n}_{\rm ran}$ from that given in equation 8 of Tadros & Efstathiou (1996), where $\bar{n}_{\rm ran}$ is the number density of unclustered points). The power spectra of the survey window function, shown in Figure 3.2, are much steeper than the expected galaxy power spectrum, falling off as $\propto k^{-4}$ for wavenumbers $k > 0.06h\,{\rm Mpc}^{-1}$. Therefore the main effect of the convolution with the survey window function is to alter the shape of the power spectrum only for wavenumbers $k < 0.06h\,{\rm Mpc}^{-1}$.

Following Tadros & Efstathiou, we define a quantity with a mean value of zero:

$$\delta(\mathbf{k}) = \hat{n}_{\rm o}(\mathbf{k}) - \alpha\hat{W}_e(\mathbf{k}), \tag{3.5}$$

where α is the ratio of the number of galaxies to random points in volume limited samples, or the ratio of the sum of the weights, given by equation 3.1, for galaxies and random points in flux limited samples. The power spectrum of galaxy clustering is then estimated using:

$$\begin{aligned} P_e(k) &= \left[\frac{1}{S_{\rm ran}^2}\frac{1}{V}\sum_{k'}\left(V^2\left|W_e(k')\right|^2 - \frac{1}{S_{\rm ran}}\right)\right]^{-1} \\ &\times \left(\frac{V^2|\delta(k)|^2}{S_{\rm gal}^2} - \frac{1}{S_{\rm gal}} - \frac{1}{S_{\rm ran}}\right), \end{aligned} \tag{3.6}$$

where we have used the notation $S_{\rm gal} = \sum_{i=1}^{N_{\rm gal}} w_{\rm gal}^2$ and $S_{\rm ran} = \sum_{i=1}^{N_{\rm ran}} w_{\rm ran}^2$. In the case of a volume limited sample $S_{\rm gal} = N_{\rm gal}$, the number of galaxies in the sample, and $S_{\rm ran} = N_{\rm ran}$, the number of unclustered points used to define the survey window function.

The power spectra are computed by embedding the Durham/UKST volume into a larger cubical volume, V. The density field is typically binned onto a 256^3 mesh using nearest gridpoint assignment (we discuss the effects of aliasing and box size in Section 3.3). The Fourier transform is performed with a FFT.

3.2.4 Error analysis

We estimate the errors on the recovered power spectrum by constructing mock catalogues that have the same radial and angular selection as the Durham/UKST Survey and which have approximately the same clustering amplitude.

We extract mock Durham/UKST catalogues from the τCDM $z = 0$ *Hubble Volume* simulation (for a discussion of this simulation see Chapter 2). The simulation covers a volume of $8 \times 10^9\, h^{-3}\,\mathrm{Mpc}^3$ and thus contains roughly 10,000 independent Durham/UKST Surveys volume limited to $z_{\mathrm{max}} = 0.06$. This allows a wide range of clustering environments to be sampled, allowing a good assessment of the size of the cosmic variance for the Durham/UKST Survey.

The power spectrum of the Hubble Volume simulation is a variant of the standard CDM model known as τCDM. The shape of the power spectrum can be described by the parameter Γ, which is set to the value $\Gamma = 0.21$ for τCDM, compared with the standard CDM case where $\Gamma = \Omega h = 0.5$. (The power spectrum used in this Hubble Volume simulation follows the definition of Γ used by Efstathiou, Bond & White (1992).) This change to the power spectrum could be achieved by postulating a massive particle whose decay produces an additional contribution to the radiation density of the universe, delaying the epoch of matter radiation equality (White, Gelmini & Silk 1995). The *rms* density fluctuations in the simulation are set to be roughly consistent with the local abundance of hot X-ray clusters (White, Efstathiou & Frenk 1993, Eke, Cole & Frenk 1996). The variance in the mass contained within spheres of radius $8\, h^{-1}\,\mathrm{Mpc}$ is $\sigma_8 = 0.6$. This is smaller than found for the galaxies in the APM Galaxy Survey, where $\sigma_8^{\mathrm{gal}} = 0.84 - 0.96$ (Baugh & Efstathiou 1993, Maddox, Efstathiou & Sutherland 1996).

In order to make an accurate assessment of the errors in our recovered power spectrum we make mock catalogues in which the clustering matches as closely as possible that in the Durham/UKST Survey. To extract such catalogues from the τCDM Hubble Volume, we apply a simple biasing prescription to the density field. We first bin the density field onto a cubical grid of cell size $5\, h^{-1}\,\mathrm{Mpc}$, using a nearest gridpoint assignment scheme. We then associate a probability to each grid cell, which depends on the ratio of the cell density to the mean density, for selecting a mass particle from that cell to be a biased or 'galaxy' particle. The form of the probability that we adopt is the same as model 2 of Cole et al. (1998) (although these authors apply a Gaussian filter to smooth the density field - we have chosen the size of the cubical grid cell to roughly match the effective volume of

the Gaussian filter):

$$P(\nu) = \begin{cases} \exp(\alpha\nu + \beta\nu^{3/2}) & \text{if } \nu \geq 0, \\ \exp(\alpha\nu) & \text{otherwise,} \end{cases} \tag{3.7}$$

where ν is the number of standard deviations of the cell density from the mean cell density. We generate a random number between 0 and 1 and if the value is less than $P(\nu)$, the particle is a biased particle. We set $\alpha = 1.26$ and $\beta = -0.45$ which are similar to the values found in Cole et al. (1998). These values of α and β bias the particles in a τCDM simulation, with similar parameters to the τCDM Hubble Volume simulation used here, to match the real space APM correlation function of Baugh (1996). This method of biasing is also used in Chapter 5.

The power spectrum of the biased set of particles is shown by the dotted line in Figure 3.3, which agrees well with the amplitude of the power spectrum of APM galaxies (Baugh & Efstathiou 1993, Gaztañaga & Baugh 1998). The dashed line in Figure 3.3 shows the power spectrum of the biased points when redshift space distortions are also included in the positions of the galaxies. As expected the power is increased on large scales and damped on small scales (Kaiser 1987).

However, to accurately give an indication of the cosmic variance expected in the Durham /UKST survey, the mock catalogues have to have the same angular and radial selection function and also contain the same number of biased particles as the Durham/UKST sample under consideration. We randomly pick the position of an observer in the simulation and select all the particles that lie inside the geometry of the Durham/UKST Survey, using the periodicity of the simulation if the observer was chosen to lie close to the edge of the simulation box. The radial selection function is then imprinted on the distribution of the particles using the method described in Section 3.2.2. The biased catalogues are then randomly sparse sampled until each mock catalogue contains the required number of mock galaxies. This is repeated 40 times for each sample from the Durham/UKST survey.

The errors on the Durham/UKST Survey power spectrum are taken to be the same size as the fractional errors on the mock catalogue power spectra. This is a valid assumption when either the contribution of shot noise to the power spectrum is negligible or, as in this case by design, the mock catalogue power spectrum and the Durham/UKST power spectrum have similar shapes and amplitudes. The errors obtained from the mock catalogues converge when averaged over 40 mock observers and are in reasonable agreement with the size of the errors obtained using the expression given in Equation 2.4.6 of

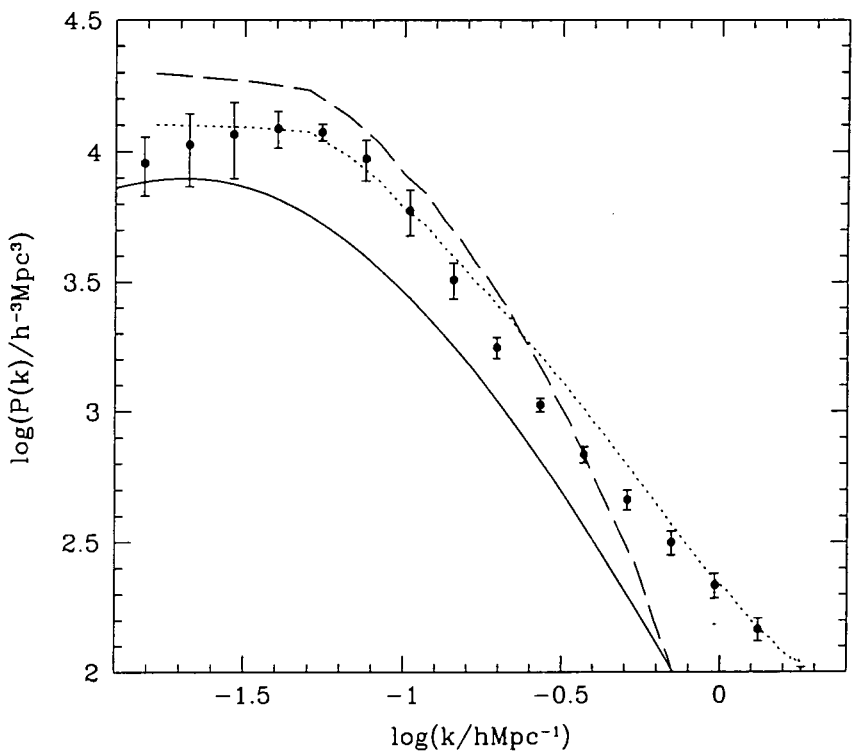

Figure 3.3: The solid line shows the linear power spectrum of the mass in the $z = 0$ τCDM Hubble Volume simulation. The dotted line shows the power spectrum of a subset of the particles in the simulation, selected according to the biasing prescription outlined in Section 3.2.4, measured in a cubical volume of side $375\,h^{-1}$ Mpc. The dashed line shows the power spectrum of these biased particles when the density is binned using redshift space coordinates. The points show the power spectrum of APM Survey galaxies, measured in real space.

Feldman, Kaiser & Peacock (1994).

3.3 Tests of the power spectrum estimation

In this Section, we make systematic tests of the power spectrum estimator (equation 3.6) in order to assess the range of wavenumbers over which we can make a robust measurement of the true power spectrum of the galaxy clustering.

On large scales, there are two main effects that can cause the recovered power spectrum to differ from the true power spectrum. First, Figure 3.2 shows that the assumption that the power spectrum of the survey window function is sharply peaked does not hold for wavenumbers $k \leq 0.04\,h\,\mathrm{Mpc}^{-1}$. On these scales, the recovered power spectrum has a different shape to the underlying power spectrum; the convolution of the power spectrum of the survey window function with the true galaxy power spectrum alters both the shape and amplitude of the estimated power spectrum at these wavenumbers. Second, the number of galaxies used in equation 3.6 is estimated from the sample itself. If fluctuations in galaxy density exist on the scale of the survey, this number can be sensitive to the environment sampled by the mock catalogue, and hence can be different from the true mean galaxy density, which is obtained by considering a much larger volume. This leads to an underestimation of the power on large scales (Peacock & Nicholson 1991, Tadros & Efstathiou 1996) which is sometimes called the integral constraint. In addition, there will be a contribution to this effect from Poisson sampling noise, even in the absence of clustering on the scale of the survey.

In Figure 3.4 we show various redshift space power spectra estimated from sets of biased tracers of the mass distribution in the $z = 0$ τCDM *Hubble Volume* simulation. The solid line shows the average power spectrum obtained from 40 cubical volumes with side $375\,h^{-1}\,\mathrm{Mpc}$. As the volume is cubical, the power spectrum is free from any effects of survey geometry. The dashed line shows the convolution of the mean power spectrum, estimated from the cubical boxes, with the window function of the survey. To convolve this power spectrum with the window function of the Durham/UKST Survey, we use equation 2.1.6 in Feldman, Kaiser & Peacock (1994) but for simplicity, we just integrate over k rather than k_x, k_y, k_z. The open circles show the power spectrum averaged over 40 mock Durham/UKST catalogues. The errorbars show the 1σ variance over these 40 mock catalogues. There are still density fluctuations over volumes the size of the Durham/UKST Survey, which leads to a variance in the number of galaxies between

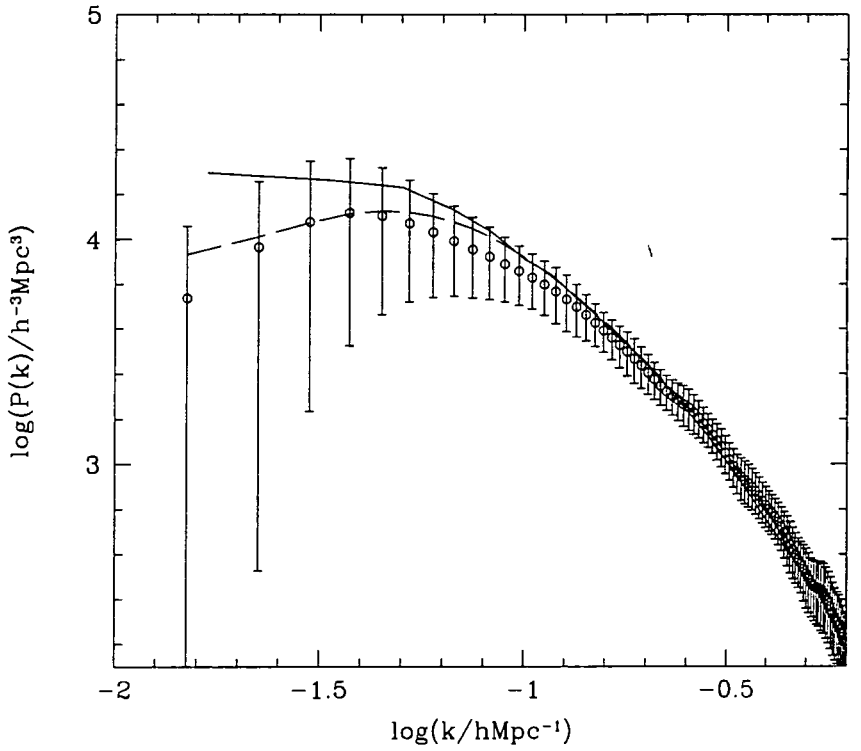

Figure 3.4: The solid line shows the redshift space power spectrum for biased particles from the $z = 0$ τCDM Hubble Volume simulation, averaged over 40 cubical volumes of side $375\,h^{-1}\,\mathrm{Mpc}$. The open circles show the power spectrum averaged over 40 mock Durham/UKST catalogues, to a volume limit of $z_{\mathrm{max}} = 0.06$. The errorbars on these points are the 1σ errors for a single power spectrum extracted from the Durham/UKST survey. The dashed line shows the convolution of the mean power spectrum measured from the large cubical volumes (solid line) with the window function of the survey.

different mock observers and this causes a bias in the power spectrum estimate at large scales.

Figure 3.4 shows that the dominant effect on the shape of the power spectrum on large scales, is the window function convolution rather than the integral constraint for the Durham/UKST Survey. The convolution with the window function power spectrum introduces curvature into the recovered power spectrum at wavenumbers, $k \lesssim 0.05\,h\,\mathrm{Mpc}^{-1}$. The real turnover in the τCDM power spectrum occurs at $k \sim 0.02\,h\,\mathrm{Mpc}^{-1}$.

The Fourier transform of the galaxy density field is computed by binning the galaxy density field onto a finite grid and then performing a Fast Fourier Transform (FFT). This can lead to spurious features in the power spectrum or aliasing of power on scales around the Nyquist frequency of the FFT grid. The magnitude of this effect is also sensitive to the scheme used to assign galaxies to the density grid. Figure 3.5 shows a series of tests designed to show the scales at which aliasing can distort the shape of the recovered power spectrum. In Figure 3.5(a), we vary the dimension of the FFT grid within a fixed box size of $1600\,h^{-1}\,\mathrm{Mpc}$, whilst in Figure 3.5(b), we vary the size of the box in which the mock catalogue is embedded for the FFT, and keep the dimension of the FFT grid fixed at 256^3. Figure 3.5(b) shows that using a 256^3 FFT grid and a box size of $800\,h^{-1}\,\mathrm{Mpc}$, gives accurate results down to $k \sim 0.6\,h\,\mathrm{Mpc}^{-1}$ or $10\,h^{-1}\,\mathrm{Mpc}$.

As we cannot infer the true mean density of galaxies from the single observed realisation of the galaxy distribution, or equivalently, we do not know the shape of the true power spectrum on these scales, we do not attempt to correct the power spectrum at large scales for either the 'integral constraint' bias or for the convolution with the power spectrum of the survey window function. Instead, the tests in this section demonstrate that the estimates of the power spectrum for the Durham/UKST Survey should be a robust measurement of the true galaxy power spectrum over the wavenumber range $0.05\,h\,\mathrm{Mpc}^{-1} \leq k \leq 0.63\,h\,\mathrm{Mpc}^{-1}$; this corresponds to a range of scales of $120\,h^{-1}\,\mathrm{Mpc}$ to $10\,h^{-1}\,\mathrm{Mpc}$; the latter is roughly the mean separation of galaxies in a volume limited sample.

3.4 Results

In this Section, we analyse volume limited and flux limited samples drawn from the Durham/UKST Galaxy Redshift Survey. In all cases, the power spectra are computed by embedding the survey in a box of side $840\,h^{-1}\,\mathrm{Mpc}$ and binning the galaxies on a grid

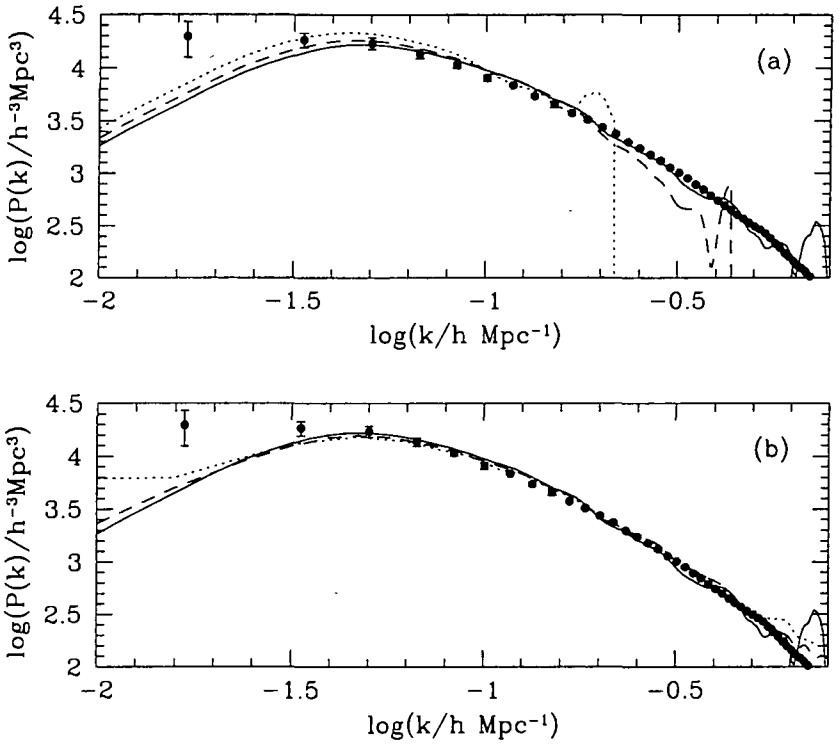

Figure 3.5: The points in (a) and (b) show the redshift space power spectrum of biased particles averaged over 40 large cubical boxes extracted from the $z = 0$ τCDM Hubble Volume simulation. (a) shows the effects of changing the size of the FFT grid when the mock catalogue is embedded in a fixed size box of side $1600h^{-1}$ Mpc . The solid line shows the result when the density grid has 256 cells per side, the dashed line has 128 cells and the dotted line has 64 cells. (b) shows the effects of varying the size of the transform box at a fixed FFT grid size of 256 cells per side. The solid line shows the results for a transform box of $1600h^{-1}$ Mpc, the dashed line for $800\,h^{-1}$ Mpc and the dotted line for $400\,h^{-1}$ Mpc.

of 256 cells on a side. We have rebinned the estimated power spectrum in bins of width $\delta k = 0.015\,h\,\mathrm{Mpc}^{-1}$ in order to reduce the correlations between the estimated power in adjacent wavenumbers. Selected results are given in Table 3.2.

The power spectra of different volume limited samples of the Durham/UKST Survey are shown in Figure 3.6. The errorbars are computed using the fractional variance in the power averaged over mock catalogues extracted from the $z = 0$ τCDM Hubble Volume simulation. Mock catalogues were made for each volume limit. As discussed in Section 3.2.4, these catalogues satisfy the same selection criteria and have approximately the same clustering as the Durham/UKST Survey galaxies. Varying the maximum redshift used to define the volume limited catalogue has two effects on the properties of the extracted sample. Increasing z_{max} increases the depth of the sample, thereby allowing fluctuations on larger scales to be probed. At the same time, however, the corresponding absolute magnitude limit imposed on the galaxies selected becomes brighter. This means that the population of galaxies used to map out the clustering varies and it is possible that intrinsically brighter galaxies could be more strongly clustered than fainter galaxies (Park et al. 1994, Loveday et al. 1995). There is a shift in the amplitude of the power spectrum as larger values of z_{max} are considered. However, the power spectra of the different samples are all consistent within the 1σ errors.

The clustering in the flux limited Durham/UKST Survey is shown in Figure 3.7. Again, the errorbars show the 1σ errors obtained from the fractional variance over the power estimated from mock catalogues made with the same selection criteria. The different panels are for weight functions (equation 3.1) using a range of constant values for the power spectrum, as indicated in the legend on each panel. Increasing the value of the power used in the weight, causes the weight function to rise at progressively larger distances (see Figure 3 of Feldman, Kaiser & Peacock 1994). This means that the effective volume probed increases and thus the sensitivity to longer wavelength fluctuations increases.

If there are no systematic problems with the survey, changing the value of the power used in the weight function defined by equation 3.1 should have little effect upon the amplitude of the recovered power spectrum (see the power spectrum analysis of the combined 1.2Jy and QDOT surveys by Tadros & Efstathiou (1995)). However, the size of the errors on a particular scale will change, depending upon whether or not the choice of weight function used really is the minimum variance estimator for the amplitude of power at these scales.

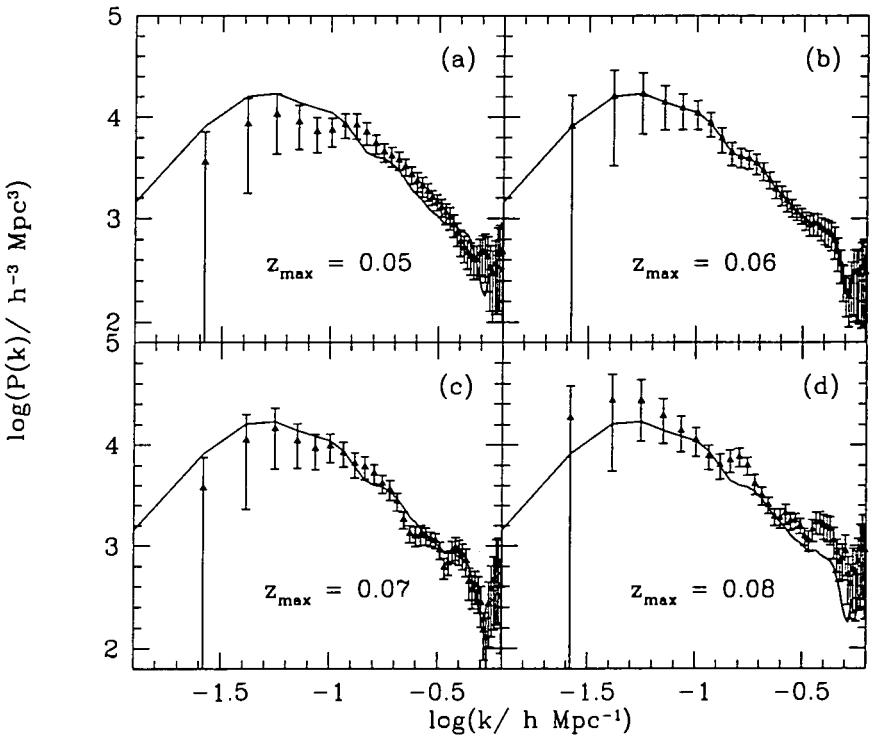

Figure 3.6: The power spectrum of the Durham/UKST Survey for different volume limited samples. The error bars are the 1σ variance obtained from the fractional errors on the power found in mock catalogues with the same angular and radial selection and approximately the same clustering. The power spectra are estimated using a box of side $840\,h^{-1}$ Mpc and a 256^3 FFT grid. The solid line is the mean power for a volume limit defined by $z_{\rm max} = 0.06$, the sample that contains the most galaxies, and is reproduced in each panel.

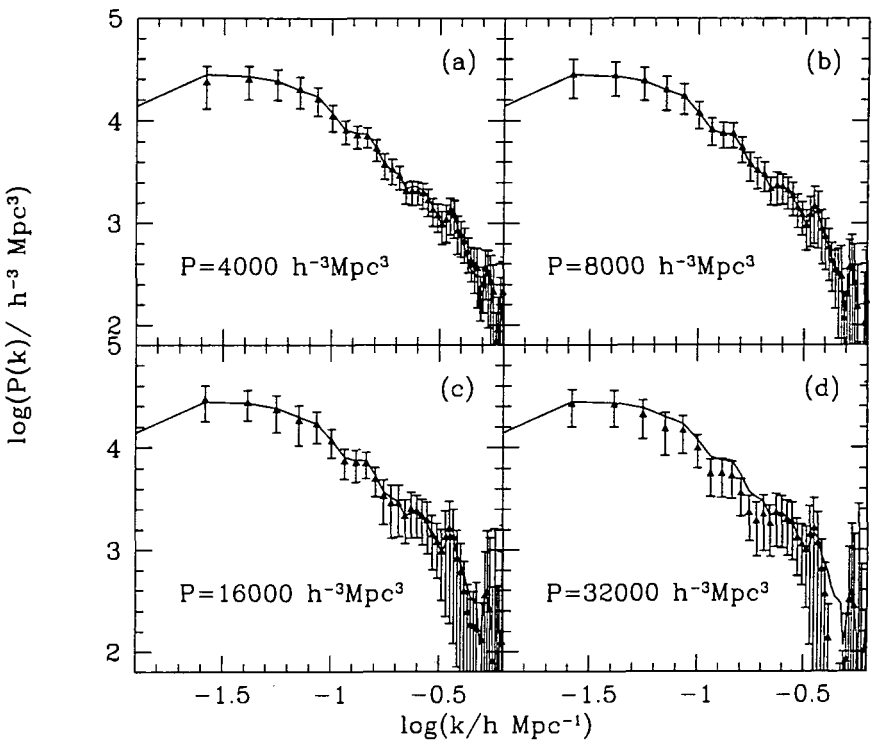

Figure 3.7: The power spectrum of the flux limited Durham/UKST Survey, for different constant values of $P(k)$ used in the weight function given in equation 3.1. The values of $P(k)$ used are 4000, 8000, 16000 and 32000 $h^{-3}\,\mathrm{Mpc}^3$ as marked in the panels. The errorbars show the 1σ variance obtained from mock catalogues with the same selection and similar clustering. The solid line is the power spectrum for a weight with $P(k) = 8000\,h^{-3}\,\mathrm{Mpc}^3$ and is reproduced in all the panels. The power spectra are computed in a box of side $840\,h^{-1}\,\mathrm{Mpc}$ using a 256^3 density grid.

The line that is reproduced in each panel of Figure 3.7 shows the power estimated for a weight function with $P(k) = 8000\,h^{-3}\mathrm{Mpc}^3$. This reference line shows that there is a negligible change in the mean power when $P(k)$ is varied by a factor of eight over the range $P(k) = 4000$–$32000\,h^{-3}\mathrm{Mpc}^3$. The flux limited power spectrum with a weight using $P(k) = 4000\,h^{-3}\mathrm{Mpc}^3$ has the smallest errorbars over the range of wavenumbers plotted, though the errors are not significantly larger for the other estimates of the power spectrum. The errors on the power spectrum measured from the volume limited sample with $z_{\mathrm{max}} = 0.06$ are larger than the errors on the power spectrum obtained from the flux limited sample for wavenumbers $k < 0.1\,h\,\mathrm{Mpc}^{-1}$; however, for wavenumbers $k > 0.1\,h\,\mathrm{Mpc}^{-1}$ the power spectrum of the volume limited sample has smaller errors.

The comparison between the power spectra of the flux limited and volume limited samples is difficult to interpret. Neither the volume nor the way in which the galaxies are weighted can be simply related between the two methods for constructing galaxy samples. Furthermore, volume limited samples select intrinsically brighter galaxies as the volume is increased and it is possible that these galaxies could have different clustering properties compared with fainter galaxies. Nevertheless, the agreement between the power spectra measured from the flux and volume limited samples is very good; if we compare the power spectrum from the volume limited sample with $z_{\mathrm{max}} = 0.06$, which contains the most galaxies, and the power spectrum with the smallest errors from the flux limited survey (i.e. with a value of $P(k) = 4000\,h^{-3}\mathrm{Mpc}^3$ used in the weight function), we find they agree within the 1σ errors. This is a further argument against a significant dependence of clustering strength upon intrinsic luminosity within the survey.

3.5 Comparison with other measurements of the power spectrum

We compare the results with measurements of the power spectrum made from other surveys, described in Chapter 2, in Figures 3.8, 3.9 and 3.10. In Figure 3.8, we compare the power spectrum from a sample of the Durham/UKST Survey, defined by a volume limit of $z_{\mathrm{max}} = 0.06$ (filled circles) with the power spectrum of a sample drawn from the Stromlo-APM Survey (Tadros & Efstathiou 1996) with the same selection (open circles). The two estimates of the power spectrum are in remarkably good agreement, except near wavenumbers of $\log(k/h\mathrm{Mpc}^{-1}) = -0.8$, -0.5 and -0.3, where there are sharp dips in the Stromlo-APM power spectrum. The solid line shows the real space power spectrum

measured from the APM Survey (Baugh & Efstathiou 1993), which is below the power spectra measured from the redshift surveys.

We compare estimates of the power spectrum made from flux limited samples in Figure 3.9. Again, the filled circles show the power spectrum of the Durham/UKST Survey, the open circles show the Stromlo-APM Survey and the crosses show the power spectrum measured from the Las Campanas Survey (Lin et al. 1996). The Durham/UKST and Stromlo-APM Surveys have similar magnitude limits, $b_J \sim 17$, whereas the Las Campanas Survey is approximately 1–1.5 magnitudes deeper, going to an $R-$band magnitude of 17.3 – 17.7, depending upon the spectrograph used to measure redshifts in a particular field. The Las Campanas survey consists of six $1.5° \times 80°$ strips and an attempt has been made to deconvolve the survey window function to give the estimate of the power spectrum plotted here (Lin et al. 1996). The power spectra from flux limited samples are in good agreement down to a wavenumber of $\log(k/h\mathrm{Mpc}^{-1}) = -1.1$ or for scales $r < 80\,h^{-1}\,\mathrm{Mpc}$. On larger scales than this, the power spectrum measured from the Las Campanas Survey is below that obtained from the Durham/UKST and Stromlo-APM Surveys, which continue to rise to $r \sim 150\,h^{-1}\,\mathrm{Mpc}$. On scales larger than this, the convolution with the survey window function of these surveys affects the shape of the recovered power spectrum. Note that the weighting scheme used to estimate the Las Campanas power spectrum is different to that employed in this paper, with each galaxy weighted by the inverse of the selection function.

In Figure 3.10, we compare the power spectrum of the Durham/UKST Survey, which is an optically selected sample, with the power spectrum obtained from an analysis by Tadros & Efstathiou (1995) of the combined 1.2Jy Survey (Fisher et al. 1995) and QDOT Survey (Efstathiou et al. 1990) datasets, which are selected in the infrared from the IRAS point source catalogue. We have plotted the minimum variance estimate of the power spectrum obtained for each dataset. The filled circles show the Durham/UKST power spectrum and the open circles show the power spectrum of IRAS galaxies. The IRAS galaxy power spectrum has a lower amplitude than the Durham/UKST power spectrum. The solid line shows the result of multiplying the IRAS power spectrum points by a constant, relative bias factor squared of $b^2_{\mathrm{IRAS}} = 1.7$, which agrees with the value inferred by Peacock & Dodds (1994).

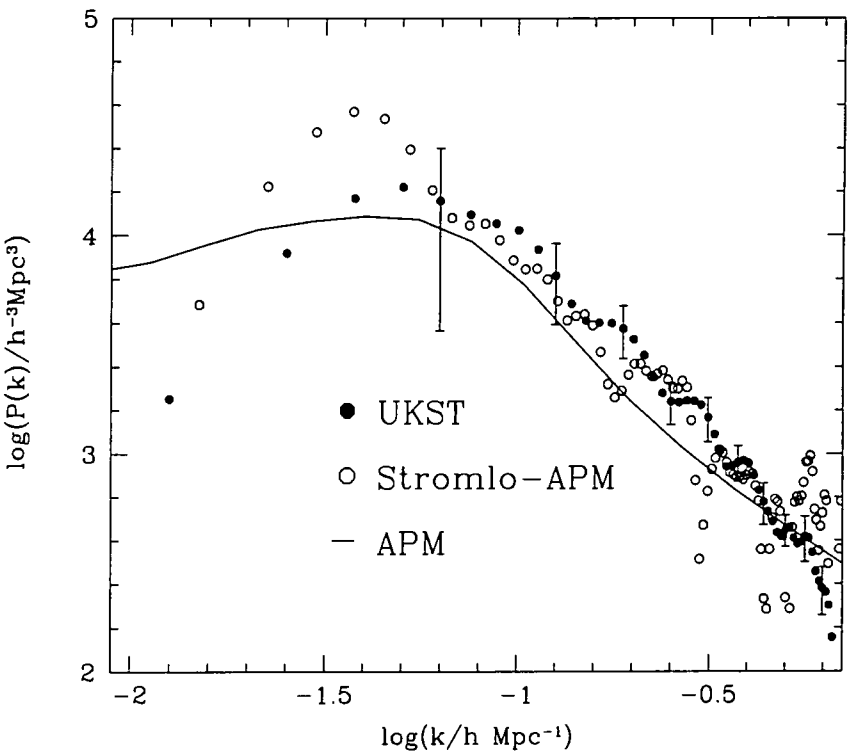

Figure 3.8: The volume limited power spectrum of the Durham/UKST Survey (solid points) compared with the power spectrum of the Stromlo-APM Survey (Tadros & Efstathiou 1996); in both cases, the volume limit is defined by $z_{\max} = 0.06$. The solid line shows the real-space APM galaxy power spectrum from Baugh & Efstathiou (1993).

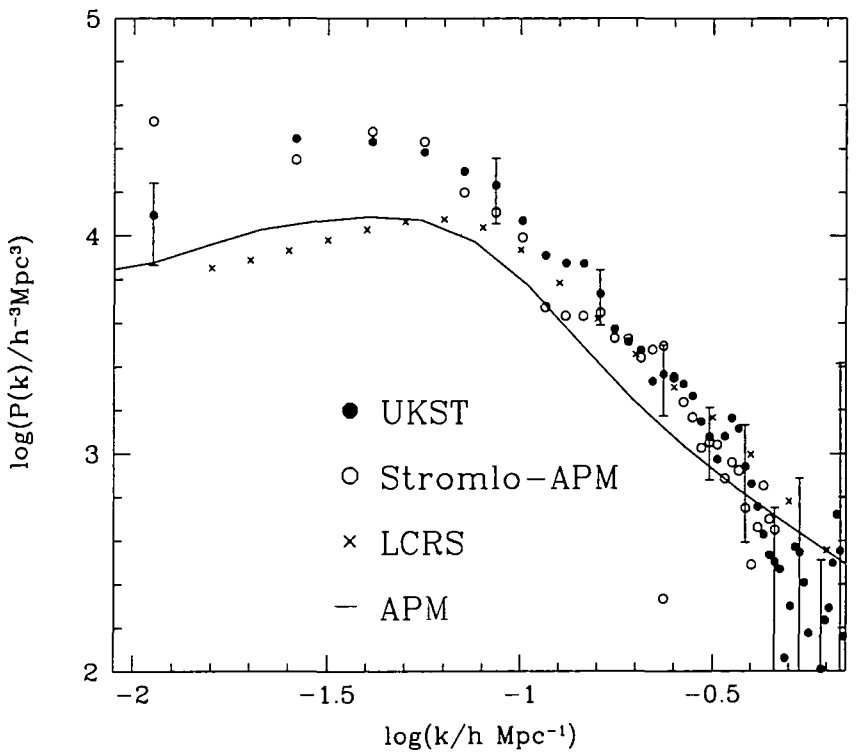

Figure 3.9: The flux limited, $P = 8000\,h^{-3}\,\mathrm{Mpc}^3$, power spectrum of the Durham/UKST Survey (solid points) compared with the flux limited power spectra of other optical samples. The open circles show the power spectrum of the Stromlo-APM Survey (Tadros & Efstathiou 1996), again flux limited with $P = 8000\,h^{-3}\,\mathrm{Mpc}^3$ and the crosses show the deconvolved $P(k)$ from the Las Campanas Redshift Survey from Lin et al. (1996).

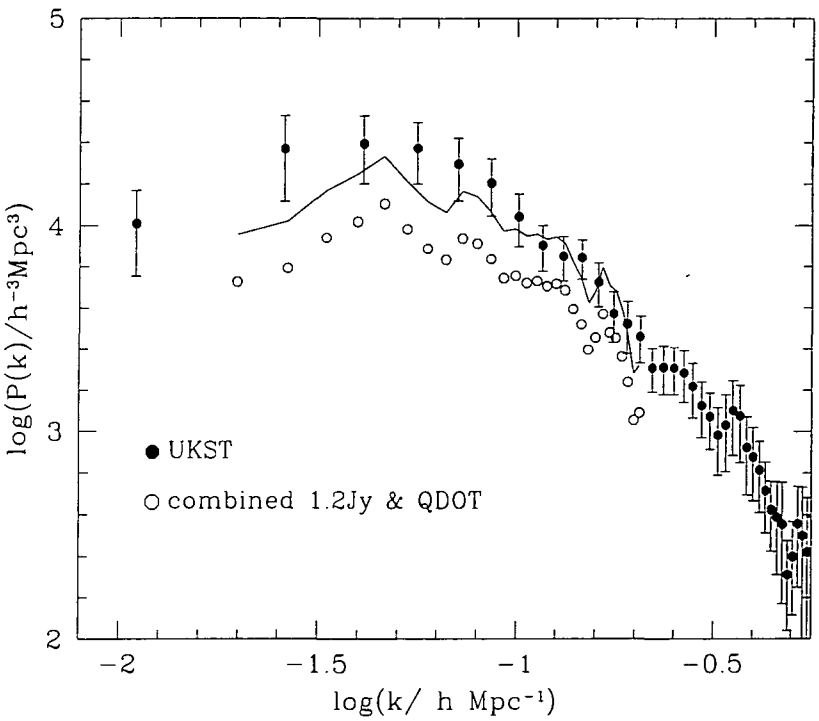

Figure 3.10: The power spectrum of Durham/UKST galaxies (filled circles) compared with the power spectrum of IRAS galaxies (open circles) obtained from the combined 1.2Jy and QDOT Surveys by (Tadros & Efstathiou 1995). Both power spectra are measured from flux limited samples and are minimum variance estimates for the respective surveys. The line shows the IRAS power spectrum after multiplying by a relative bias factor of $b^2_{\rm IRAS} = 1.7$, where we have assumed that the bias factor is not a function of scale.

3.6 Implications for models of Large Scale Structure

In this Section we compare the predictions of various scenarios for the formation of large scale structure in the Universe with the power spectrum of the Durham/UKST Survey.

There are several steps that one has to go through in order to compare a power spectrum for the mass distribution, calculated in linear perturbation theory, with a galaxy power spectrum measured using the positions of the galaxies inferred from their redshifts:

(i) Compute the non-linear power spectrum of the mass distribution given the amplitude of *rms* density fluctuations specified by the value of σ_8. We use the transformation given by Peacock & Dodds (1996).

(ii) Choose a bias parameter, b, relating fluctuations in the mass distribution to fluctuations in the galaxy distribution: $P_{\rm gal}(k) = b^2 P_{\rm mass}(k)$. In the following analysis we make the simplifying assumption that the bias parameter is independent of scale.

(iii) Model the distortion of clustering due to the fact that galaxy redshifts have a contribution from motions introduced by inhomogeneities in the local gravitational field as well as from the Hubble flow.

(iv) Convolve the power spectrum with the window function of the Durham/UKST survey.

On large scales, (iii) leads to a boost in the amplitude of the power spectrum (Kaiser 1987), whilst on small scales the power is damped by random motions inside virialised groups and clusters. It is important to model these two extremes and the transition between them accurately, as this can have a significant effect on the shape of the power spectrum over the range of scales considered. We model the effects of the peculiar motions of galaxies on the measured power spectrum using the formula given by Peacock & Dodds (1994):

$$P_s(k) = b^2 P_r(k) G(\beta, y) \tag{3.8}$$

where $P_s(k)$ is the galaxy power spectrum measured in redshift space and $P_r(k)$ is the mass power spectrum measured in real space. The function $G(\beta, y)$, where $\beta = \Omega^{0.6}/b$, the factor that relates particle velocities to the rate of growth of the density field, see Appendix B, and $y = k\sigma$ (σ is the one dimensional velocity dispersion), is given by:-

$$G(\beta, y) = \frac{\sqrt{\pi}}{8} \frac{\mathrm{erf}(y)}{y^5} (3\beta^2 + 4\beta y^2 + 4y^4)$$

$$-\frac{\exp[-y^2]}{4y^4}(\beta^2(3+2y^2)+4\beta y^2). \quad (3.9)$$

This assumes that the small scale peculiar velocities of galaxies are independent of separation and have a Gaussian distribution. To convolve the power spectra with the window function of the Durham/UKST Survey, we use equation 2.1.6 in Feldman, Kaiser & Peacock (1994). Rather than integrate over k_x, k_y, k_z we just integrate over k, this approximation is seen to be accurate in Figure 3.4 as the convolved power spectrum (shown by the dashed line) matches closely the mock Durham/UKST power spectrum (open circles).

We compare the models with the Durham/UKST power spectrum measured from a sample with a volume limit defined by $z_{\rm max} = 0.06$. This power spectrum measurement has slightly larger errors than the minimum variance power spectrum from the flux limited sample on large scales, $r \sim 60\,h^{-1}\,$Mpc. However, on scales smaller than this, the volume limited power spectrum has the smallest errors.

We test the simple model for the transformation of a linear theory power spectrum for mass fluctuations to a galaxy power spectrum measured in redshift space in Figure 3.11. The open circles show the mean power spectrum from 10 Durham/UKST mock catalogues, using the real space coordinates of the particles to map out the density. The filled circles show the distortion caused to the power spectrum when the peculiar motions of the particles are included. The lines show the results of applying equation 3.8 to the linear theory τCDM power spectrum. This equation results from performing an azimuthal average over the angle between the line of sight and the wavevector of the density fluctuation. This assumption will mainly affect the longest wavelength fluctuations in a real survey that does not cover the whole sky. These scales are already distorted by the convolution with the survey window function. The model provides a reasonably good fit for a one dimensional velocity dispersion of $\sigma = 500{\rm km\,s^{-1}}$, which is approximately the value found in the GIF simulations (Jenkins et al. 1998).

The first test we perform is to compare the power spectrum of APM Survey galaxies (Baugh & Efstathiou 1993, Baugh & Efstathiou 1994a, Gaztañaga & Baugh 1998) with the Durham/UKST volume limited power spectrum. The APM power spectrum is measured in real space and is estimated by inverting the angular correlation function of APM galaxies with $17 \leq b_J \leq 20$. The shapes of the real space and redshift space power spectra can be compared in Figure 3.12(a). The real space power spectrum is shown by the dashed line, after multiplying by a constant factor of 1.4 to match the Durham/UKST Survey at small wavenumbers, so that the relative shapes of the real space and redshift space

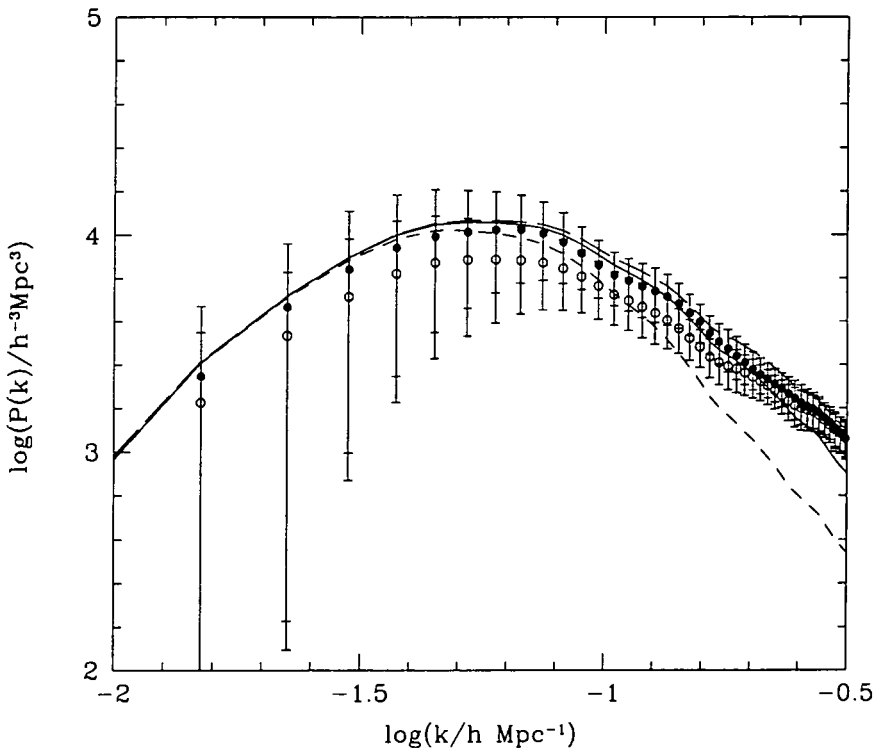

Figure 3.11: The open (filled) circles show the mean power averaged over 10 mock Durham/UKST Surveys in real (redshift) space. The lines show the Peacock and Dodds predictions (equation 3.8) with a bias of $b = 1.5$ and $\sigma = 300\,\mathrm{km\,s^{-1}}$ (long dashed), $\sigma = 500\,\mathrm{km\,s^{-1}}$ (solid) and $\sigma = 1000\,\mathrm{km\,s^{-1}}$ (short dashed).

power spectra can be readily compared. We have rebinned the Durham/UKST power spectrum and error bars to match the binning of the APM power spectrum which has $\delta \log(k/h\mathrm{Mpc}^{-1}) = 0.13$. The spacing of the power spectrum measurements is now much larger than the half width of the survey window function, so there is essentially no covariance between the errors at different wavenumbers. The rebinned Durham/UKST power spectrum is shown in each panel of Figure 3.12 by the points and errorbars. We retain this binning of the Durham/UKST power spectrum in the subsequent analysis of theoretical power spectra below. As we are comparing two galaxy power spectra, the factor of b^2 is omitted in equation 3.8. The best fitting APM galaxy power spectrum over the range of scales $10 < r < 100h^{-1}$Mpc, including the redshift space distortions, is shown by the solid line in Figure 3.12(a). The transformation into redshift space removes the inflection in the real space APM power spectrum around a wavenumber of $k \sim 0.15\,h\,\mathrm{Mpc}^{-1}$. The best fitting values of β and σ, with 1σ errors are $\beta = 0.60 \pm 0.35$ and $\sigma = 320 \pm 140\,\mathrm{kms}^{-1}$. Tadros & Efstathiou (1996) found $\beta = 0.38 \pm 0.67$ by comparing the Stromlo-APM redshift space power spectrum to the APM Survey power spectrum, restricting their attention to wavenumbers in the range $0.05 < k < 0.1\,h\,\mathrm{Mpc}^{-1}$, over which, they argued that the damping of power in redshift space is negligible. The one dimensional velocity dispersion that we recover from the comparison is in excellent agreement with the measurement of Ratcliffe et al. (1998c), but has much larger errors. By considering the galaxy correlation function binned in separation parallel and perpendicular to the line of sight, Ratcliffe et al. (1998c) obtained a value for the pairwise *rms* velocity dispersion along the line of sight of $\sigma_{||} = 416 \pm 36\,\mathrm{kms}^{-1}$. This quantity is approximately $\sqrt{2}$ times the one dimensional velocity dispersion that we find, giving $\sigma = 294 \pm 25\,\mathrm{kms}^{-1}$. If we add in quadrature the estimated error in the measured redshifts $\sim 150\,\mathrm{kms}^{-1}$ (Ratcliffe et al. 1998d), the Ratcliffe et al measurement implies $\sigma = 330\mathrm{kms}^{-1}$.

We also test the popular Cold Dark Matter (CDM) models by treating the bias parameter and the one dimensional velocity dispersion as free parameters. The mass power spectra are found using the transfer function given in Efstathiou, Bond & White (1992). The models that we consider are; $\Omega_\mathrm{m} = 1$ CDM with a shape parameter $\Gamma = 0.5$ and with a normalisation of $\sigma_8 = 0.52$ (SCDM) that reproduces the local abundance of rich clusters (Eke, Cole & Frenk 1996); a model with a normalisation of $\sigma_8 = 1.24$ and $\Gamma = 0.5$ (COBE-SCDM), which matches the COBE detection of temperature anisotropies in the microwave background, but seriously over-predicts the abundance of hot clusters; τCDM, with $\Omega_\mathrm{m} = 1$, $\Gamma = 0.2$ and $\sigma_8 = 0.52$, which simultaneously matches the amplitude implied

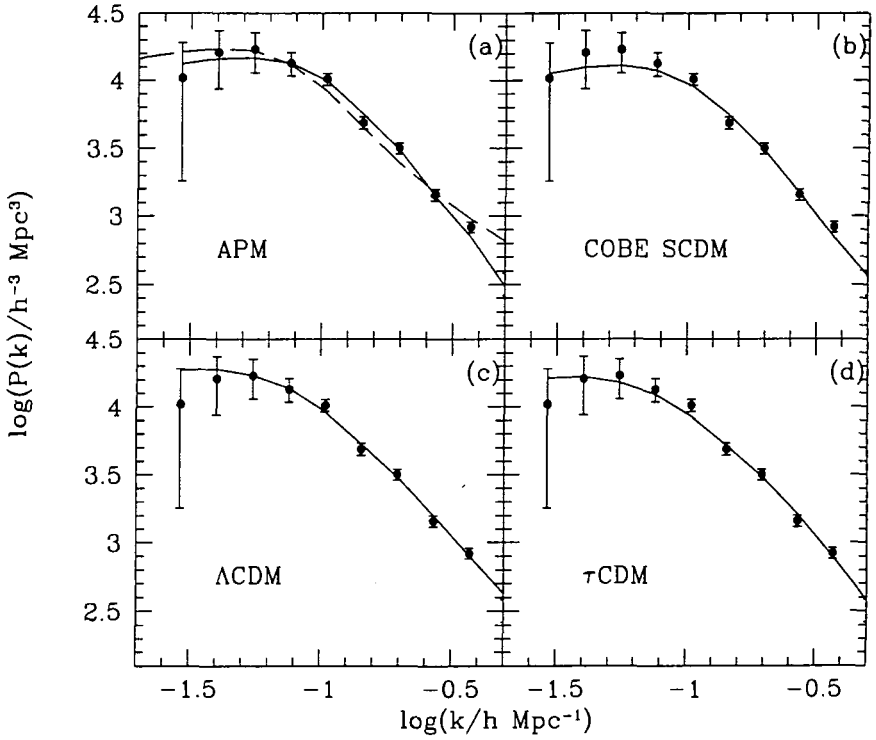

Figure 3.12: The points in each panel show the Durham/UKST power spectrum for a volume limited sample with $z_{\max} = 0.06$. The power spectrum estimates have been rebinned to reduce the covariance in the errors. In (a), the dashed line shows the APM galaxy power spectrum measured in real space, rescaled to match the Durham/UKST power spectrum at large scales. The solid line shows the APM power spectrum, including the effects of distortion in redshift space. The remaining panels, b, c, d, show the best fitting curves for several variants of the Cold Dark Matter model. Table 3.6. gives the values of the linear bias b and the one dimensional velocity dispersion σ used in equation 3.9.

Model	σ_8	Γ	h	Ω_m	b	β	σ (km s^{-1})
SCDM	0.52	0.5	0.50	1.0	2.97±0.26	0.34	1240±200
COBE-SCDM	1.24	0.5	0.50	1.0	0.91±0.1	1.10	760±130
τCDM	0.52	0.2	0.50	1.0	1.64±0.17	0.61	320±110
ΛCDM	0.93	0.2	0.67	0.3	1.04 ±0.09	0.47	520±100

Table 3.1: The parameters of each of the CDM models and the best fitting values of the bias parameter, b, and the one dimensional velocity dispersion, σ, for various different cosmological models. We also give the value of β implied by the best estimate of the bias parameter, b, and the density parameter Ω_m of the model.

by COBE and by the cluster abundance through an adjustment to the shape of the power spectrum, as described in Section 3.2.4, and ΛCDM, which is a low density model, with a present day value for the density parameter of $\Omega_m = 0.3$ and a cosmological constant of $(\Lambda_0 c^2)/(3H_o^2) = 0.7$ (Efstathiou, Sutherland & Maddox 1990). The ΛCDM model has a normalisation of $\sigma_8 = 0.93$, matching the amplitude implied by both COBE and the abundance of clusters.

The best fitting parameters are given in Table 3.6. Note that as we specify a value for the density parameter, Ω_m, through the choice of structure formation model, we are constraining the value of the bias parameter b; the implied errors on β are much smaller than if we had not selected a value for Ω_m beforehand. For all the models considered, reasonable agreement with the Durham/UKST Survey power spectrum can be obtained if no restrictions are placed on the values of the bias and one dimensional velocity dispersion that are used in the fit. However, the SCDM and COBE-CDM models only produce a reasonable fit to the Durham/UKST power spectrum if large values of the velocity dispersion are adopted; these values are inconsistent with the value we obtain from the comparison with the APM Survey power spectrum at more than 3σ. The velocity dispersion required for the ΛCDM model is marginally inconsistent (1.5σ) with the value that we infer from the comparison with the APM power spectrum. This agrees with the results of the complementary analysis of the two point correlation function carried out by Ratcliffe et al. (1998b), who analysed the clustering in a N-body simulation with a very similar cosmology and power spectrum. The τCDM model gives the best fit to the Durham/UKST data in the sense that the values of β and σ required are in excellent

agreement with those obtained from the comparison with the real space galaxy power spectrum.

3.7 Conclusions

There is remarkably good agreement between measurements of the power spectrum of galaxy clustering made from optically selected surveys, on scales up to $80\,h^{-1}\,\mathrm{Mpc}$. For scales larger than $80\,h^{-1}\,Mpc$, we find good agreement between the power spectra of the Durham/UKST Survey and of the Stromlo-APM Survey (Tadros & Efstathiou 1996). We measure more power on these scales than is found in a clustering analysis of the Las Campanas Redshift Survey (Lin et al. 1996). We find no convincing evidence for a dependence of galaxy clustering on intrinsic luminosity within the Durham/UKST Survey. However, we do measure a higher amplitude for the power spectrum from our optically selected sample compared with that recovered for galaxies selected by emission in the infrared; the offset in amplitude can be described by an optical/infrared bias factor squared of $b^2_{\mathrm{IRAS}} = 1.7$.

We have compared the shape and amplitude of the APM Survey power spectrum (Baugh & Efstathiou 1993, Baugh & Efstathiou 1994a, Gaztañaga & Baugh 1998), which is free from any distortions caused by peculiar velocities, with the Durham/UKST power spectrum. The APM power spectrum displays an inflection at $k \sim 0.15\,h\,\mathrm{Mpc}^{-1}$. Using a simple model for the effects of galaxy peculiar velocities that is valid over a wide range of scales, we find that the inflection is removed in redshift space. The APM power spectrum can be distorted to give a good match to the Durham/UKST power spectrum for $\beta = \Omega_{\mathrm{m}}^{0.6}/b = 0.60 \pm 0.35$ and a one dimensional velocity dispersion of $\sigma = 320 \pm 140\,\mathrm{kms}^{-1}$. These values are consistent with those found from an independent analysis of clustering in the Durham/UKST Survey by Ratcliffe et al. (1998c), who obtained $\beta = 0.52 \pm 0.39$ (see Hamilton (1998) and references therein for estimates of β made from different surveys using a range of techniques) and $v_{12}(\sim \sqrt{2}\sigma) = 416 \pm 36\,\mathrm{kms}^{-1}$. The value of β that we obtain from this analysis, can be used, with an assumption for the value of Ω_{m}, to infer the amplitude of fluctuations in the underlying mass distribution. For example, if we assume $\Omega_{\mathrm{m}} = 1$, the value for β suggests that APM galaxies are biased with respect to fluctuations in the mass by $b = 1.7 \pm 1.0$; this in turn implies a value for the *rms* fluctuations in mass of $\sigma_8 = 0.84/b = 0.50 \pm 0.29$, which is consistent with that required to reproduce the abundance of massive clusters (Eke, Cole & Frenk 1996). As the abundance of clusters

and β have a similar dependence on $\Omega_{\rm m}$, this agreement will hold for any value of $\Omega_{\rm m}$ and therefore does not constrain $\Omega_{\rm m}$.

We have compared theoretical models for structure formation with the power spectrum of the Durham/UKST survey. The best agreement is found with a variant of the CDM model known as τCDM. A low density model with a cosmological constant also provides reasonable agreement, but for a velocity dispersion that is marginally inconsistent with that obtained from the comparison between the power spectra of the Durham/UKST and APM Surveys. Critical density CDM models with shape parameter $\Gamma = 0.5$ require one dimensional velocity dispersions that are much too high to provide a good fit to the Durham/UKST power spectrum. One possible way to resolve this problem would be to relax the assumption that the bias parameter between galaxies and the mass distribution is independent of scale. Whilst a constant bias is undoubtedly a poor approximation on small scales (e.g. Coles 1993, Mann, Peacock & Heavens 1998, Benson et al. 2000b), this analysis probes scales greater than $20\,h^{-1}$ Mpc. A scale dependent bias on such large scales could be motivated in a cooperative galaxy formation picture (Bower et al. 1993), though the higher order moments of the galaxy distribution expected in such a model are not favoured by current measurements (Fry & Gaztañaga 1994).

$kh\mathrm{Mpc}^{-1}$	$P(k)_{vol,z_{\max}0.06}$	1σ	$P(k)_{flux,P=4000}$	1σ	$P(k)_{flux,P=8000}$	1σ
0.0411	16153	12867	24828	8946	26962	9936
0.0561	17014	10170	23596	7805	24177	8556
0.0711	13927	6467	19763	6602	19775	7243
0.0860	12215	4672	15992	4880	17049	5684
0.1000	11096	3530	11332	3262	12091	3606
0.1078	9974	2962	9645	2635	10103	2960
0.1161	8682	2363	7991	2011	8113	2353
0.1251	7137	1973	7446	1836	7732	2190
0.1348	5707	1571	7065	1667	7481	2065
0.1452	4546	1168	6986	1515	7464	2009
0.1565	4153	934	5782	1339	6012	1670
0.1686	3937	825	4477	1149	4547	1351
0.1817	3713	798	3566	998	3553	1150
0.1958	3305	727	3171	871	3170	1051
0.2110	2750	549	2582	649	2696	854
0.2273	2189	413	2039	511	2230	733
0.2449	1790	371	2034	530	2289	779
0.2639	1517	339	1934	540	2112	722
0.2844	1247	234	1572	470	1727	626
0.3064	1091	184	1221	371	1251	456
0.3302	932	191	997	372	1028	474
0.3558	907	224	1262	498	1451	827
0.3834	826	209	882	360	929	514
0.4131	745	211	667	251	594	314
0.4451	615	213	421	156	344	199
0.4796	331	119	311	175	239	237
0.5168	181	87	339	170	338	279
0.5568	305	196	239	199	206	304
0.6000	339	252	111	115	40	89

Table 3.2: Measurements of the power spectrum from the Durham/UKST Survey. The first column gives the wavenumber; logarithmically spaced bins are used for $k > 0.1\,h\,\mathrm{Mpc}^{-1}$. The second column gives the power measured in a volume limited sample with $z_{\max} = 0.06$. The fourth and sixth columns give the power measured in the flux limited Durham/UKST Survey, when weights of $P = 4000\,h^{-3}\,\mathrm{Mpc}^3$ and $P = 8000\,h^{-3}\,\mathrm{Mpc}^3$, respectively are used in equation 3.1. Columns 3, 5 and 7 gives the 1σ errors on each measurement. The errors are the 1-σ variance from 40 mock catalogues extracted from the $z = 0$ τCDM Hubble Volume.

Chapter 4

Higher Order Clustering in Galaxy Surveys

4.1 Introduction

Maps of the local universe have improved dramatically over the last decade and permit the clustering pattern of galaxies to be quantified on large scales (e.g. Efstathiou et al. 1990, Maddox et al. 1990b, Saunders et al. 1991). Such observations can potentially constrain both the nature of the dark matter and the statistics of primordial density fluctuations.

The first accurate measurements of the galaxy two-point correlation function on scales greater than $10h^{-1}$Mpc indicated more structure than expected in the simplest form of the cold dark matter (CDM) model (Efstathiou, Sutherland & Maddox 1990). This led to variants of the CDM model being studied. Currently, the most successful CDM model is a low density, spatially flat universe with a cosmological constant, ΛCDM. The power spectrum in the ΛCDM model is described by a shape parameter $\Gamma \sim 0.2$. If fluctuations in the dark matter are normalised so as to reproduce the local abundance of hot X-ray clusters (White, Efstathiou & Frenk 1993), the power spectrum in the ΛCDM model is similar to that observed for galaxies on scales around $k \sim 0.05 - 0.2h\mathrm{Mpc}^{-1}$ (Gaztañaga & Baugh 1998). On small scales, however, when the effects of peculiar velocities are ignored (real space), the dark matter power spectrum has a higher amplitude than the galaxy power spectrum (e.g. Gaztañaga 1995, Peacock 1997, Jenkins et al. 1998).

Heuristic biasing schemes, in which the galaxy distribution is proposed to be a local transformation of the smoothed density field, have enjoyed a certain degree of success in reproducing the observed correlation function (e.g. Coles 1993, Cole et al. 1998, Mann, Peacock & Heavens 1998, Narayanan et al. 1999). Progress towards a physical understand-

ing of the processes responsible for producing a bias between the galaxy and dark matter distributions has been made using semi-analytic models for galaxy formation (Kauffmann et al. 1999, Benson et al. 2000b). In a ΛCDM model that reproduces the bright end of the field galaxy luminosity function, Benson et al. find remarkably good agreement with both the amplitude and power law slope of the correlation function of the real space APM Survey galaxies (Baugh 1996) and the redshift space correlation function from the ESO Slice Project (Guzzo et al. 2000).

The constraints on models of galaxy formation provided by the two-point correlation function are somewhat limited. The second moment gives a full statistical description of the density field only in the case of very weak fluctuations. Galaxy clustering can be described in more detail if the J-point, volume-averaged, correlation functions, $\bar{\xi}_J$, are extracted. If the clustering results from the gravitational amplification of a Gaussian primordial density field, then the J-point functions are predicted to follow a hierarchical scaling, $\bar{\xi}_J = S_J \bar{\xi}_2^{J-1}$. The amplitudes S_J do vary with scale, but at a much slower rate than the volume-averaged correlation functions (Juszkiewicz et al. 1993). This scaling behaviour has been studied extensively for CDM in N-body simulations (e.g. Bouchet, Schaeffer & Davis 1991, Baugh, Gaztañaga & Efstathiou 1995, Gaztañaga & Baugh 1995, Hivon et al. 1995, Colombi, Bouchet & Hernquist 1996, Szapudi et al. 1999b)

Fry & Gaztañaga (1993) proposed a simple bias model, based on the assumption that fluctuations in the galaxy distribution can be written as a function of the dark matter fluctuations, when both fields are smoothed on large scales where $\bar{\xi}_2 \ll 1$. The model gives predictions for the moments of the galaxy distribution in terms of the moments for the dark matter. To leading order in the dark matter variance, the galaxy variance is given by $\bar{\xi}_2^{\rm gal} = b^2 \bar{\xi}^{\rm DM}$, where b is usually called the linear bias. To the same order, an additional or second order bias factor, b_2, is required to specify the galaxy skewness:

$$S_3^{\rm gal} = \frac{1}{b}(S_3^{\rm DM} + 3\frac{b_2}{b}). \tag{4.1}$$

Gaztañaga & Frieman (1994) discuss the implications of the measurements of S_J from the APM Survey for the bias parameters in this model.

In this Chapter, we analyse the clustering in two optically selected redshift surveys that sample large volumes of the local universe. The Durham/UKST Survey (Ratcliffe et al. 1998d) and Stromlo-APM Survey (Loveday et al. 1996), discussed in Chapter 2, are magnitude limited to $b_J \sim 17$. Galaxies are sparsely sampled from the parent catalogues at a rate of 1-in-3 in the case of the Durham/UKST Survey and 1-in-20 for the Stromlo-

APM Survey. The Stromlo-APM Survey covers a three times larger solid angle than the Durham/UKST Survey. By combining the results from the two surveys, S_J can be determined over a large dynamic range in cell size.

4.2 Counts in Cells Methodology

The method of counts in cells has been used theoretically and observationally to study clustering many times in the literature (see for example Peebles 1980, Gaztañaga 1992, Baugh et al. 1995). The standard method is to place a large number, M, of cells of radius l within the survey volume and to count the number of galaxies that lie within each cell. The Jth central moment is then given by

$$m_J(l) = \frac{1}{M}\sum_{i=1}^{M}(N_i - \bar{N})^J, \tag{4.2}$$

The discreteness of the particles means a shot noise term needs to be subtracted from the moments, this is particularly important on small scales where the expected count in each cell is ~ 1. To correct for this contribution a Poisson shot-noise model is applied to the moments. The moments up to $J = 4$ are given below:-

$$m_2^{sn}(l) = m_2 - \bar{N}, \tag{4.3}$$

$$m_3^{sn}(l) = m_3 - 3m_2 + 2\bar{N}, \tag{4.4}$$

$$m_4^{sn}(l) = m_4 - 3m_2^2 - 6m_3 + 11m_2 - 6\bar{N}. \tag{4.5}$$

The volume averaged J-point correlation function is given by

$$\bar{\xi}_J(l) = \frac{m_J^{sn}(l)}{\bar{N}^J}, \tag{4.6}$$

and the definition of the higher order moments is then

$$S_J = \frac{\bar{\xi}_J}{\bar{\xi}_2^{J-1}}, \tag{4.7}$$

This method has been adapted by Szapudi & Szalay (1993). The basis of their method is to throw a large number of cells onto the galaxy distribution in order to obtain the probability distribution of finding N galaxies in a cell of a given size l.

$$P_N = \frac{1}{M}\sum_{i}^{M}\delta(N_i = N) \tag{4.8}$$

where again M is the number of cells and N_i is the number of objects within the cell. The moments of the count probability distribution are estimated using the factorial moment

technique, which automatically adjusts the moments to compensate for the sampling of a continuous density field using discrete galaxies (Szapudi & Szalay 1993; Szapudi, Meiksin & Nichol 1996). This approach is analogous to the method outlined above.

The approach that we adopt to measure the moments of the galaxy count probability distribution differs in two respects from most previous work. A similar methodology is applied to the PSCz Survey by Szapudi et al. (2000). The first difference lies in how the higher order moments are extracted from the redshift survey. The count probability distribution is measured in a series of volume limited samples of varying radial depth drawn from the flux limited survey. The moments obtained for a particular cell volume are compared between the different volume limited samples and the minimum variance estimate is adopted as our measurement for this scale. The construction of volume limited samples is straightforward: a maximum redshift for the sample is defined and any galaxy from the flux limited redshift survey that would remain visible if displaced out to this redshift is included in the sample (see Chapter 3 for a description of how to construct a volume limited sample).

The number density of galaxies in a volume limited sample is independent of radial distance. This is in direct contrast to a flux limited survey, where the number density changes rapidly with radius, as seen in Figure 2.3 in Chapter 2. To analyse the count distribution in a flux limited catalogue, a weight must be assigned to each galaxy to compensate for the radial selection function. The analysis of volume limited samples is therefore much simpler as the need to devise an optimal weighting scheme and to construct a suitable estimator of the moments is avoided. The results in any case can be shown to be equivalent (Colombi, Szapudi & Szalay 1998).

This approach does, however, rely upon the assumption that galaxy clustering does not depend on luminosity, at least over the range of luminosities that we consider in our samples (see column 4 of Table 4.1 for the absolute magnitudes that define the volume limited samples we analyse). Loveday et al. (1995) measured the two-point correlation function in redshift space for galaxies selected from the Stromlo-APM survey on the basis of absolute magnitude. These authors found no significant evidence for a difference in clustering amplitude when comparing samples over a much broader range of absolute magnitudes than we consider in our analysis. Similar conclusions were reached by Tadros & Efstathiou (1996) who analysed the amplitude of the power spectrum in different volume limited samples drawn from the same survey. A weak effect, at just over the 1σ level, was seen only for the deepest sample, corresponding to an absolute magnitude of $M_{b_J} = -20.3$.

In Chapter 3, we find that the power spectra in volume limited samples drawn from the Durham/UKST survey vary by less than the 1σ errors as the depth of the sample is changed. Therefore, the approximation that the intrinsic clustering is the same in different volume limited samples is fully justified by previous work on the surveys we analyse in this Chapter.

The second difference from previous work is the treatment of the errors on the measured moments. A theoretical calculation of the errors is made using the method described by Szapudi, Colombi & Bernardeau (1999a) *. All the possible sources of statistical error are included in the calculation:

• Finite survey volume: The finite volume of the survey means that fluctuations on scales larger than the survey volume are not probed at all. In addition, fluctuations on scales approaching the maximum dimensions of the survey are poorly sampled.

• Edge effects: The density field around galaxies that lie close to the survey boundary is not sampled as well as the density field around a galaxy that is well within the boundary. This is because cells are not permitted to straddle the survey boundary.

• Discreteness: The underlying density field is assumed to be continuous. Sampling this field discretely with galaxies makes an additional contribution to the measured moments.

• Sampling or measurement errors due to the finite number of cells used to construct the count probability distribution.

The theoretical calculation of the errors requires a number of quantities to be specified beforehand. Some of these, namely the measured values of the variance and S_J for a given cell size and the sample volume, are estimated directly from the sample. The other quantities, the variance over the full sample volume and the higher order cumulant correlators, are treated as parameters. The errors that we obtain are fairly insensitive to reasonable choices for the values of these parameters (for a full discussion see Szapudi, Colombi & Bernardeau 1999a).

The theoretical error calculation has been extensively tested for clustered distributions of dark matter using N-body simulations (Colombi et al. 2000). As a further check of the error calculation, we have measured the dispersion in the moments over mock Durham/UKST Survey samples with a redshift limit of $z = 0.06$ extracted from the τCDM and ΛCDM $z = 0$ Hubble Volume N-body simulations, as described in Chapters 2 and 3. A comparison between the two different error estimation techniques is given

*The FORCE package (FORtran for Cosmic Errors) was used to compute errors. It is available upon request from its authors, S. Colombi (colombi@iap.fr) or IS (szapudi@cita.utoronto.ca).

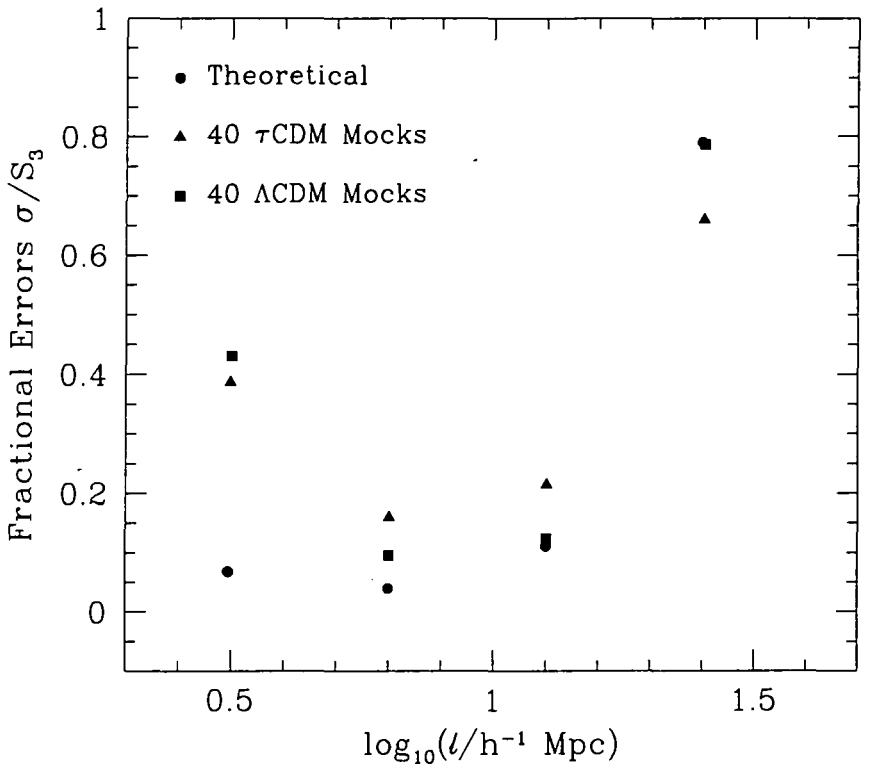

Figure 4.1: We compare the fractional errors on S_3 for two different methods: using the theoretical calculation outlined here and given in more detail in Szapudi, Colombi & Bernardeau (1999a) (circles) and using mock catalogues to obtain the errors. The triangles show the fraction errors from 40 biased τCDM mock catalogues, described in Chapter 3 and the squares show the fractional errors from 40 ΛCDM mock catalogues, which are not biased as the power spectrum from the ΛCDM mock catalogues is similar to that of the Durham/UKST Survey.

in Figure 4.1. The circles show the fraction error from the theoretical calculation given in Szapudi, Colombi & Bernardeau (1999a), the triangles show the fraction error from a sample of 40 τCDM mock catalogues and the squares show the fractional errors from a sample of 40 ΛCDM mock catalogues. Apart from on the very smallest scale, the errors from the two different estimators agree to within 20%.

4.3 Results

The galaxy count probability distribution is measured in cubical cells of side $3-40h^{-1}$Mpc in a series of volume limited samples drawn from the Durham/UKST and Stromlo-APM redshift surveys. The limiting redshifts of the samples are in the range $z \sim 0.05 - 0.08$, corresponding to maximum radial depths of $140–220h^{-1}$Mpc. The higher order moments are calculated from the count probability distribution using the factorial moment technique introduced by Szapudi & Szalay (1993). In practice, measurement errors, (iv) in the list of statistical errors given in Section 4.2, are negligible in comparison to the other contributions, because on the order of 10^8 cells are used to determine the count distribution at each scale.

The second moment or variance of the galaxy distribution is shown in Figure 4.2. In both panels, the filled circles show measurements obtained from the Durham/UKST Survey and the open circles show those from the Stromlo-APM Survey. Figure 4.2(a) shows the variance as a function of cell size in volume limited samples extracted from the survey, with a maximum redshift of $z = 0.06$. Figure 3.1 in Chapter 3 shows that the number of galaxies as a function of the maximum redshift defining a volume limited sample peaks at this redshift for both surveys. These results are in good agreement with estimates of the variance made from the surveys using different techniques. The solid line shows an independent estimate of the variance obtained from the the power spectrum of the same volume limited sample from the Durham/UKST Survey for wavenumbers $k \leq 0.43h\text{Mpc}^{-1}$. We have used the approximate transformation between the power spectrum and the variance given in Peacock (1991). The dotted lines show the 1σ error on this estimate, which comes directly from the error on the measured power spectrum. The very good level of agreement between these different estimates demonstrates that large volume cells genuinely measure fluctuations on large scales. Our results for a volume limited subsample of the Stromlo-APM survey agree well with those obtained from the full magnitude limited survey shown by the crosses in Figure 4.2(a) (Loveday et al. 1992).

Survey	Cell size (h^{-1}Mpc)	R_{max} (h^{-1}Mpc)	Volume ($10^6 h^{-3}$Mpc^3)	M_{crit} - 5log$h$	N_{gal}	S_3	S_4
Durham/UKST	3.125	170	0.721	-19.58	510	1.94±0.14	1.5
Durham/UKST	6.3125	180	0.855	-19.73	515	2.11±0.08	5.0±3.8
Durham/UKST	12.625	180	0.855	-19.73	515	1.82±0.21	3.0
Durham/UKST	25.	170	0.721	-19.58	510	1.67±1.32	2.2
Stromlo-APM	3.9375	180	2.547	-19.45	471	2.07±0.57	13.
Stromlo-APM	8.875	180	2.547	-19.45	471	1.89±0.17	3.1
Stromlo-APM	18.1875	190	2.995	-19.58	465	2.24±0.29	8.2
Stromlo-APM	36.625	200	3.493	-19.71	434	1.41±1.01	-

Table 4.1: Minimum variance estimates of S_3 and S_4 in cubical cells from the Durham/UKST and the Stromlo-APM Surveys. The errors on S_3 are the 1σ theoretical errors The relative errors on the estimates of S_4 are greater than 100% apart from for one Durham/UKST value.

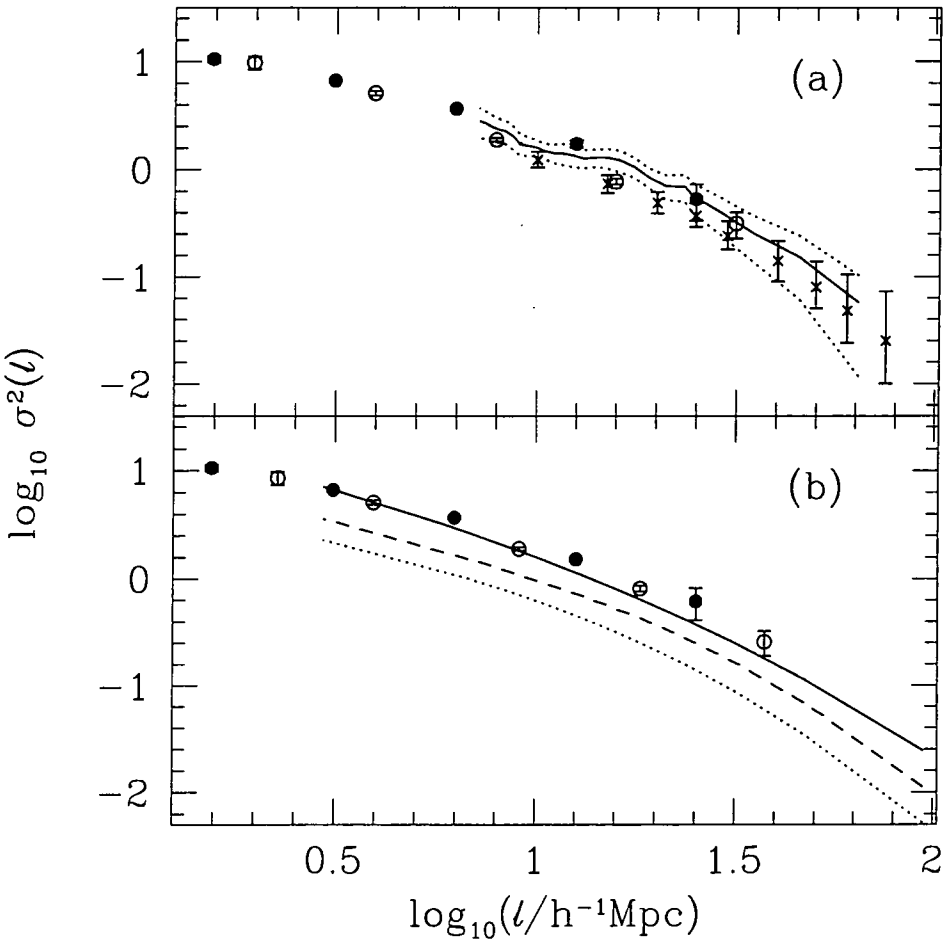

Figure 4.2: In both panels, solid circles show the variance in the Durham/UKST Survey, whilst open circles show the Stromlo-APM Survey results. In (a), we show the variance in volume limited samples with $z_{max} = 0.06$. The solid line shows an estimate of the variance made from the power spectrum measured in the same Durham/UKST sample (see Chapter 3, Figure 3.6); the dotted lines show the 1σ errors. The crosses show the variance for the flux limited Stromlo-APM survey from Loveday et al. (1992). The error bars on these points show 95% confidence limits. In (b), the circles show the best estimates of the variance, extracted from a series of volume limited samples. The lines show the variance in redshift space for the N-body simulations discussed in Section 4.4. The lines show linear power spectra described by $\Gamma = 0.2$ and $\sigma_8 = 1$ (solid), $\Gamma = 0.5$ and $\sigma_8 = 1$ (dashed) and $\Gamma = 0.5$, $\sigma_8 = 0.66$ (dotted).

The error bars on these points show the 95% percent confidence limits and are computed under the assumption that the distribution of fluctuations is Gaussian.

In Figure 4.2(b), the points show the best estimates of the variance extracted from the two surveys, as described in Section 4.1 and in Chapter 2. The best estimates of the variance from the Durham/UKST survey come from two samples, with radial limits of $R_{max} = 170\ h^{-1}$Mpc and $R_{max} = 180\ h^{-1}$Mpc; reading from left to right, the first two points and the last point in Figure 4.2(b) come from the $R_{max} = 170\ h^{-1}$Mpc sample, whilst the third and fourth points come from the $R_{max} = 180\ h^{-1}$Mpc sample. The smoothness of the locus traced out by the points supports our assumption that there is no significant dependence of clustering strength on luminosity. The lines in 4.2(b) show the variance in a set of representative CDM simulations; these simulations are discussed in Section 4.4.

The minimum variance estimates of S_3 from the Durham/UKST and Stromlo-APM surveys are listed in Table 4.1, along with the properties of the volume limited sample in which the measurement was made. The errors on S_3 are the 1σ theoretical errors predicted for a sample of this volume and geometry and containing the stated number of galaxies. For cubical cells between $3 - 20h^{-1}$Mpc, we find remarkably little variation in the value of S_3, with errors in the range $10 - 20\%$, which again provides further evidence against any significant luminosity dependence of clustering. We obtain S_3 on scales larger than $20h^{-1}$Mpc, but with much larger errors.

When the relative error on the estimate of S_J approaches 100%, the perturbative techniques used in the error calculation break down. Nevertheless, the calculation still reliably indicates that the errors are large and that the measurement has no significance. The relative errors on S_4 are estimated to be >100% on all scales in the Stromlo-APM survey. There is only one scale where S_4 can be reliably constrained from the Durham/UKST survey. This scale is also the scale on which S_3 is most accurately measured and the sample that gives the best measurement of S_4 is the sample that also gives the best measurement of S_3, $R_{max} = 170h^{-1}$Mpc. As we expect this to be the case in general, the values for S_4 from the same sample as the minimum variance measurements of S_3 are listed in Table 4.1. These estimates should be treated with caution as the errors are large.

4.4 Discussion

The mean values we obtain for the skewness are in agreement with those found in shallower redshift surveys, though we find errors that are somewhat larger (e.g. Gaztañaga 1992, Bouchet et al. 1993, Fry & Gaztañaga 1994, Benoist et al. 1999) . A comprehensive list of references can be found in Table 1 of Hui & Gaztañaga (1999). Moreover, in spite of the relatively large volumes of the surveys considered in this analysis, we find that a significant measurement of S_4 is only possible at one scale. There are two main reasons for the discrepancy in the magnitude of the estimated errors. The first is that some previous results are quoted as averages over the values of S_3 determined on different scales, exploiting the relatively flat form of S_3 in redshift space. This leads to smaller errorbars under the incorrect assumption that the individual measurements are independent. The second reason is that not all of the contributions to the statistical errors listed in Section 4.2 were considered in previous analyses.

We have constrained S_3 over a wide range of scales, extending beyond $l \sim 20h^{-1}$Mpc, where simple models for bias can be tested most cleanly. Indirect measurements of S_3 on these scales have been obtained from the IRAS 1.2Jy Redshift Survey by fitting a parametric functional form for the count probability distribution to the measured counts (Kim & Strauss 1998). The choice of function is not physically motivated and the error model used is simplistic and may underestimate the true variance (Gaztañaga, Fosalba & Elizalde 1999, Hui & Gaztañaga 1999). Szapudi et al. (2000) have measured S_3 from the IRAS PSCz survey, using the same techniques employed in this paper, and find $S_3 = 0.87 \pm 0.48$ for cells of side $l = 37h^{-1}$Mpc, which is in good agreement with the value we find, quoted in Table 4.1.

We compare our measurements of S_3 with the values inferred from the parent angular catalogues of the redshift surveys in Figure 4.3 (Gaztañaga 1994; Szapudi, Meiksin & Nichol 1996). The results from the angular catalogues are obtained by first extracting the projected count distribution on the sky, and then applying a deprojection algorithm to infer the moments in three dimensions. The algorithm requires knowledge of the survey selection function. The deprojected angular measurements are in real space and are free from any distortion due to the peculiar motions of galaxies. On large scales, the mean value we find for S_3 in redshift space is below that found in real space. However, the errors are large on both measurements, and the results are consistent at the 1σ level. Moreover, it is somewhat unclear exactly how important edge effects in the angular measurements

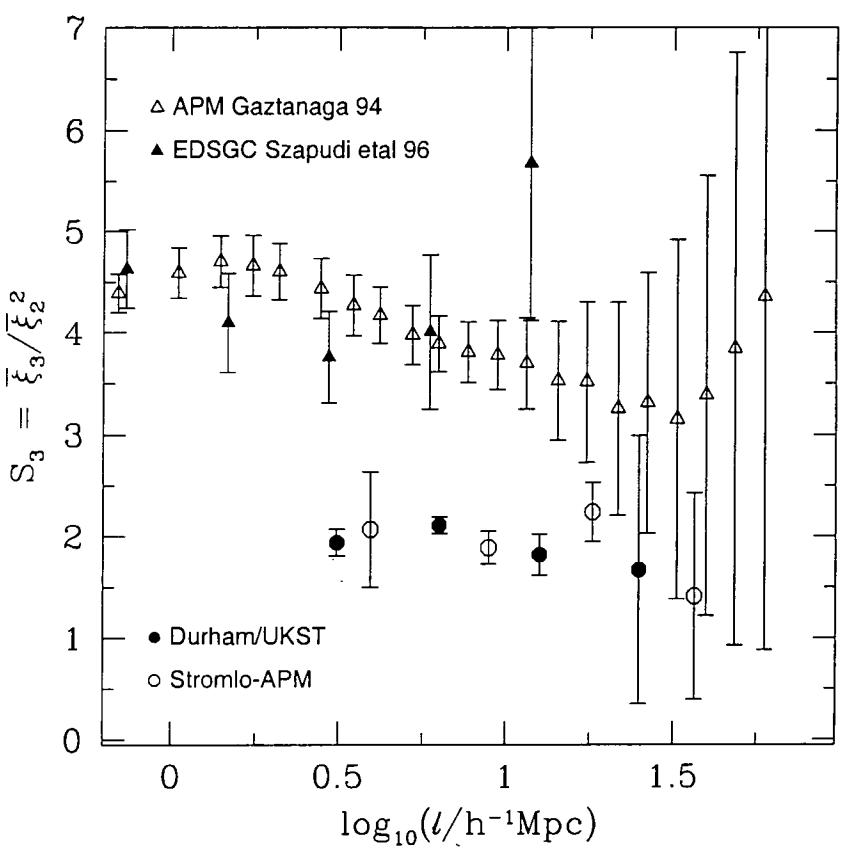

Figure 4.3: The skewness extracted from the redshift surveys (filled circles show Durham/UKST results, open circles show Stromlo-APM results) compared with the three dimensional values inferred from the parent angular catalogues (the open triangles show the APM Survey results from Gaztañaga (1994) and the filled triangles show the results from the Edinburgh-Durham Southern Galaxy Catalogue from Szapudi, Meiksin & Nichol (1996)).

and systematic effects in the deprojection technique are on these large scales (e.g. Szapudi, Meiksin & Nichol 1996, Gaztañaga & Bernardeau 1998, Szapudi & Gaztañaga 1998).

On small and intermediates scales, $l \leq 15h^{-1}$Mpc, our determinations are below those obtained from the angular catalogues. This is due to redshift space distortions. The same qualitative behaviour is seen for S_3 measured in real space and redshift space in numerical simulations of hierarchical clustering. In Figure 4.4, we compare S_3 measured in the N-body simulations used by Gaztañaga & Baugh (1995), which are representative of the behaviour in CDM models, with the redshift survey results. The heavy dashed lines in each panel show S_3 in real space, and the heavy solid lines show S_3 including the effects of the peculiar motions of the dark matter. The dotted lines show the error on the mean obtained over five realisations of the initial conditions (the box size of the simulations is $378h^{-1}$Mpc). Two different power spectra are considered: panel (a) shows a model with $\Gamma = 0.2$ and (b) and (c) show a model with $\Gamma = 0.5$ at two different epochs. On large scales, the value of S_3 depends upon the shape of the power spectrum and is in good agreement with the perturbation theory predictions, which are shown by the light lines. The value of S_3 in redshift space also depends upon the shape of the power spectrum, and is insensitive to epoch or equivalently to the amplitude of the fluctuations, as shown by Figures 4.4(b) and (c). The real and redshift space values of S_3 become consistent at $l \approx 20h^{-1}$Mpc, in good agreement with the comparison presented for the data in Figure 4.3.

We now investigate how the predictions from the simulations can be reconciled with the observations and discuss the implications for biasing. The model developed by Fry & Gaztañaga (1993) predicts a relationship between the skewness in the galaxy distribution, $S_3^{\rm gal}$, and that in the underlying dark matter, $S_3^{\rm DM}$, that is applicable on large scales and which is given by equation 4.1. The variance for the dark matter in the simulation with $\Gamma = 0.2$ and $\sigma_8 = 1$ is very close to the observed variance in galaxy counts (the solid line in Figure 4.2), indicating that a relatively small linear bias term is required; at $l \sim 20h^{-1}$Mpc, the linear bias is $b = 1.16 \pm 0.06$. Furthermore, in redshift space, the linear bias is essentially independent of scale. Thus, given the scale independence of the skewness that we measure for galaxies and which is predicted for the dark matter from the simulations, we can insert the values for $S_3^{\rm gal}$, $S_3^{\rm DM}$ and b into equation 4.1 and obtain a value for the second order bias term, b_2. At $l \sim 20h^{-1}$Mpc, a second order bias term of value $b_2 = -0.20 \pm 0.14$ is required for the skewness of the dark matter to match that seen for galaxies. For the simulation with $\Gamma = 0.5$ and σ_8=0.66, the bias term is larger (the

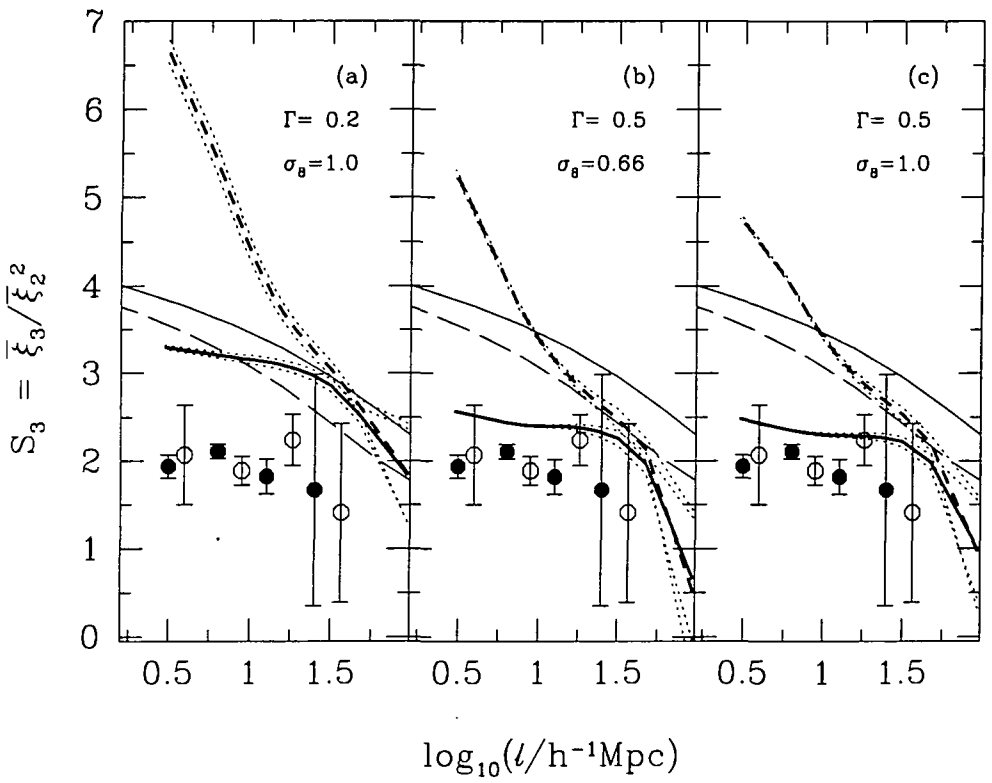

Figure 4.4: A comparison of the minimum variance measurements of the skewness listed in Table 4.1 with the skewness obtained from N-body simulations. In each panel, the filled circles show the skewness measured in the Durham/UKST Survey and the open circles show Stromlo-APM Survey results. The light lines show the linear perturbation theory predictions for S_3 in real space and are reproduced in each panel; the solid line shows the skewness for a power spectrum with $\Gamma = 0.2$, and the dashed line shows the result for $\Gamma = 0.5$. The heavy lines show the simulation results and the dotted lines show the error on the mean over five realisations of the initial density field. The heavy dashed (solid) lines show the skewness measured in real (redshift) space. The simulation outputs are described by the following sets of power spectrum parameters: (a) $\Gamma = 0.2$ and $\sigma_8 = 1$, (b) $\Gamma = 0.5$ and $\sigma_8 = 0.66$ and (c) $\Gamma = 0.5$ and $\sigma_8 = 1$.

dotted line in Figure 4.2), $b = 1.86 \pm 0.10$, and the second order bias term is $b_2 = 1.0 \pm 0.4$. Hence, whilst a linear bias term is sufficient to reconcile the variance measured in redshift space for galaxies and for dark matter, additional bias terms are required to match the results for the skewness.

A similar counts in cells analysis has been applied to the PSCz Survey, and yields values for S_3 in good agreement with those reported here (Szapudi et al. 2000). At first sight this result is intriguing, in view of the well known difference in the amplitude of the two-point functions of optical and infra-red selected galaxies on large scales (e.g. Peacock 1997, Chapter 3). Thus having demonstrated the need to consider a second order bias term in addition to the linear bias usually discussed, it would appear that both these quantities can depend on the way in which galaxies are selected. These issues are best addressed using semi-analytic models of galaxy formation (Baugh et al. 2000).

Chapter 5

Mock 2dF QSO Catalogues

5.1 Introduction

Galaxy surveys have allowed us to study the large scale structure in the local Universe, as discussed in the previous Chapters. Large clusters of galaxies and voids have been detected in these surveys and they have placed constraints on cosmological parameters and models of structure formation. However, large area galaxy redshift surveys have so far only probed clustering out to $z \lesssim 0.1$. Even the 2dF Galaxy Redshift Survey (Colless 1998) and the Sloan Digital Sky Survey (Gunn & Weinberg 1995) will only observe galaxies with redshifts $\lesssim 0.3$. Surveys such as the CNOC Survey (Yee et al. 2000) have probed galaxy clustering out to $z = 0.6$ and, in the future, the DEEP (Davis & Faber 1998) and VIRMOS (Le Fèvre 1998) Surveys will probe galaxy clustering out past $z = 1$. Steidel and collaborators have even developed techniques from which galaxies can be selected at higher redshifts. If a galaxy is observed in the B and V bands but not in the U band, known as U band drop outs or Lyman break galaxies, the galaxy must have a redshift of $z \sim 3$. This method can be extended to even higher redshifts using B band drop outs (Steidel et al. 1995, Steidel et al. 1999). However, these deep surveys currently only cover a small area of sky and clustering measurements may be prone to cosmic variance, i.e. if a different patch of sky was observed, the measured clustering may be significantly different. If we wish to study the clustering out to high redshifts over a wide area in a reasonably short time period, then different techniques or a different class of object are required.

Since the work of Osmer (1981), redshift surveys of Quasi-Stellar Objects (QSOs) have been used to probe clustering out to deeper redshifts. Originally, QSO clustering was hard to detect, due to the low number density of QSOs in the survey and the inhomogeneous nature of the surveys. However, it was demonstrated in the mid 1980's that QSOs are

clustered, at least on small $r < 10h^{-1}$Mpc scales, and that the amplitude of QSO clustering is similar to that of present day, optically selected galaxies (Shaver 1984, Shanks et al. 1986, Shanks et al. 1987). More recently, several groups have looked at the evolution of QSO clustering. Andreani & Cristiani (1992) find that the clustering amplitude of optically bright QSOs evolves slowly with redshift, such that the clustering amplitude of high redshift QSOs is 2σ higher than that of low redshift QSOs, whereas Croom & Shanks (1996) find that the amplitude of QSO clustering remains approximately constant with redshift. However, existing surveys of QSOs are still fairly small, the largest of them being the Large Bright QSO Survey (Hewitt, Foltz & Chaffee 1995) which contains 1053 QSOs. Measurements of some clustering statistics have been hampered by the small number of objects contained in the survey or the limited sky coverage of the surveys.

Our knowledge of the structure of the Universe at high redshifts will be drastically improved once the 2dF QSO Survey (Croom et al. 1998a, Boyle et al. 2000, Croom et al. 2000) and the QSO part of the Sloan Digital Sky Survey (Gunn & Weinberg 1995) are completed. These surveys will contain at least a factor of 25 more QSOs and cover a larger area of sky than existing QSO surveys. One of the key aims of both surveys is to measure the clustering of QSOs out to $\sim 1000h^{-1}$Mpc with much higher accuracy than is possible with existing surveys. For example, the increased number of QSOs in the 2dF QSO survey should reduce the errors on the measured correlation function by a factor of 5 on scales $\gtrsim 10h^{-1}$Mpc (Croom et al. 2000) and on smaller scales, the reduction of the errors maybe even larger.

In order for the clustering statistics to be measured accurately from the 2dF QSO survey, it is important that the methods of estimating these quantities and associated errors are tested thoroughly before the survey is finished. One of the most widely used methods for doing this is to construct mock catalogues that have the same geometry, sampling rate and clustering as the completed survey is expected to have (Cole et al. 1998). Different ways of measuring the correlation function, the power spectrum and other such statistics, can then be tested in advance of the completion of the survey to see which estimates give reliable results.

We construct mock catalogues from the *Hubble Volume* ΛCDM simulation (Evrard et al. 2000), where particles are output along an observer's past lightcone so the evolution of clustering in the mass is imprinted on the N-body simulation. This simulation is the largest to date, both in terms of the number of particles used (10^9) and the volume followed (27 Gpc^3). However, QSOs are thought to be biased tracers of the mass as measurements

of QSO clustering have a higher amplitude than the clustering of the dark matter particles in the *Hubble Volume* simulation. If realistic predictions of the clustering statistics from the 2dF QSO Survey are to be made, the mass particles in the *Hubble Volume* simulation must be biased to match the clustering expected from the 2dF QSO survey. We apply a simple biasing prescription, described in Cole et al. (1998), to the dark matter particles such that the mock catalogues approximately reproduce the correlation function from the first 9,000 QSOs observed in the 2dF QSO Survey. We discuss the construction of these mock catalogues in Section 5.2.

Using the mock catalogues, we test the estimators of various two-point statistics. We do not consider higher order statistics from the mock catalogues. Correctly biasing a simulation to match the correlation function and power spectrum of the 2dF QSO survey does not ensure that the higher order statistics from the simulation will match the higher order statistics that will be measured from the 2dF QSO Survey. A second order bias term may be required to reconcile statistics such as S_3 from the dark matter to S_3 from the galaxies or QSOs, as found in Chapter 4. In Section 5.3, we test if the estimation of the correlation function is sensitive to the method used and which methods of error estimation are the most reliable. We also split the mock catalogues into different redshift bins in Section 5.3.4 to test how well the evolution of QSO clustering in the 2dF QSO Survey will be measured. In Section 5.4, we test how well the power spectrum will be measured from the 2dF QSO survey and which methods allow the power spectrum to be measured over as wide a range of scales as possible. We test different error estimators and conclude by asking whether the turnover expected in CDM models will be detected. In Section 5.5 we summarize our findings.

5.2 Creating a Mock QSO Catalogue

5.2.1 The Hubble Volume Simulation

The suite of *Hubble Volume* simulations are discussed in Chapter 2 and by Evrard et al. (2000). In this Chapter, we will simply refer to the ΛCDM *Hubble Volume* lightcone output as the *Hubble Volume* simulation. The parameters of the simulation are $\Omega_{\rm m}$=0.3, Ω_Λ=0.7, $H_{\rm o}$=70 km s^{-1}Mpc^{-1} and the normalisation, σ_8, is 0.9, consistent with the abundance of hot X-ray clusters (White, Efstathiou & Frenk 1993) and with the level of anisotropies in the CMBR found by COBE (Smoot et al. 1992). The input power spectrum was calculated using CMBFAST (Seljak & Zaldarriaga 1996) with the parameters listed above

but assuming that Ω_b=0.04 and Ω_{CDM}=0.26. This changes the shape of the input power spectrum to $\Gamma = \Omega_m h \exp(-\Omega_b[\sqrt{2h} + \Omega_m]/\Omega_m)$=0.17 (Sugiyama 1995) as compared to $\Gamma = \Omega_m h = 0.21$. Throughout this Chapter, we will assume Ω_m=0.3, Ω_Λ=0.7 to be the cosmology, unless otherwise stated.

One billion mass particles are contained within a cube that is $3{,}000h^{-1}$Mpc on a side. One of the vertices was chosen to be the observer and the long axis of the lightcone was oriented along the maximal diagonal. The lightcone therefore extends to a depth of $\sim 5{,}000h^{-1}$Mpc, which corresponds to $z \sim 4$ in the ΛCDM cosmology. The solid angle of the lightcone is 75×15 degrees which is split into three 75×5 degree slices. The 2dF QSO Survey consists of 2 such slices. Ideally we would like many more than 3 slices but due to the large volume of the survey ($4.7\times10^9 h^{-3}$Mpc^3 for $z < 2.4$ and Ω_m=0.3, Ω_Λ=0.7) this is currently not possible.

5.2.2 The Clustering Amplitude

To create realistic mock catalogues, we must bias the mass particles. We could use models in the literature, such as those of Mo & White (1996), Matarrese et al. (1997) or Colín et al. (1999), to describe the QSO mass bias as a function of redshift. However, our primary aim here is to test which estimators of the clustering and errors will give reliable results when the clustering is measured from the 2dF QSO Survey. We just need to create mock catalogues that have a similar clustering pattern, and the same angular and radial selection function and sampling expected in the final 2dF QSO Survey. We therefore adopt a simpler approach and bias the mass particles in the *Hubble Volume* simulation to match the clustering found in initial measurements of the correlation function from the 2dF QSO survey.

By January 2000, ~19,000 objects from the 2dF QSO input catalogue had been observed, ~9,000 are found to be QSOs. The measurement of the two-point correlation function from these QSOs is consistent with a power law that has the form $\xi = (r/r_0)^{-\gamma}$ with $r_0 \sim 4.60\pm0.4h^{-1}$Mpc and $\gamma = 1.57\pm0.14$ over the range of scales $1 \lesssim r \lesssim 50h^{-1}$Mpc if $\Omega_m = 1$ is the assumed cosmology (Croom et al. 2000). The correlation function has been estimated in different ways to check that it is robust and not sensitive to the current incompleteness. For example, the correlation function has been measured from QSOs in areas that have 100% observational coverage as well as from all of the QSOs so far observed and the two correlation functions are found to be consistent. The values of r_0 and γ are also consistent with previous measurements of the QSO correlation function. Croom

& Shanks (1996) looked at the clustering properties of 3 different UVX selected QSO catalogues, the LBQS (Hewitt, Foltz & Chaffee 1995), the Durham/AAT survey (Boyle 1986) and the ESO/AAT+CFHT combined sample (Boyle, Jones & Shanks 1991), discussed in Chapter 2. The clustering of the whole sample was found to be consistent with a power law with an amplitude of $r_0 \sim 4 \pm 1.14h^{-1}$Mpc and a value of $\gamma = 1.4 \pm 0.4$, measured assuming the same $\Omega_m = 1$ cosmology. The correlation function from the January 2000 catalogue gives an indication of the improved signal to noise in the 2dF QSO Survey, already the errors on r_0 have already decreased from ± 1.14 to ± 0.40.

The cosmology of the *Hubble Volume* simulation is $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$. Assuming this cosmology, the correlation function of the Jan 2000 catalogue has the power law form $r_0 \sim 6.58 \pm 0.52h^{-1}$Mpc and $\gamma = 1.74 \pm 0.14$ (Croom et al. 2000). The value of r_0 is increased as the QSO-QSO separations are larger when the cosmology is $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$ as opposed to Ω_m=1. This is the value that the correlation function measured from the mock catalogues should match.

We assume that the clustering amplitude is the same at all redshifts. The evidence for this comes from Croom & Shanks (1996). They split the data from the three completed surveys mentioned above into two redshift bins, $0.3 < z < 1.4$ and $1.4 < z < 2.2$. The clustering of the QSOs measured in the two bins, assuming Ω_m=1 for the cosmology, was found to be consistent, although each redshift bin only contained around 500 QSOs. The cosmology adopted here is $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$ rather than Ω_m=1 so the clustering amplitude may not evolve in the same way as when Ω_m=1 but we will assume that the clustering is constant with redshift. Initial measurements of the clustering from the January 2000 catalogue show little evolution if ether Ω_m=1 or Ω_m=0.3, Ω_Λ=0.7 is assumed.

This assumption is also consistent with the strong clustering measured from Lyman break galaxies at redshifts $z \sim 3$ (Adelberger et al. 1998). A value of $r_0 \sim 6h^{-1}$Mpc, assuming $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$ is inferred from the counts-in-cells measurements of these galaxies.

5.2.3 The Biasing Prescription

The biasing prescription used in this Chapter is similar to the method described in Chapter 3, Section 3.2.4 and follows method 2 as described by Cole et al. (1998). The bias probability is based on the density field at the epoch at which particles are selected rather

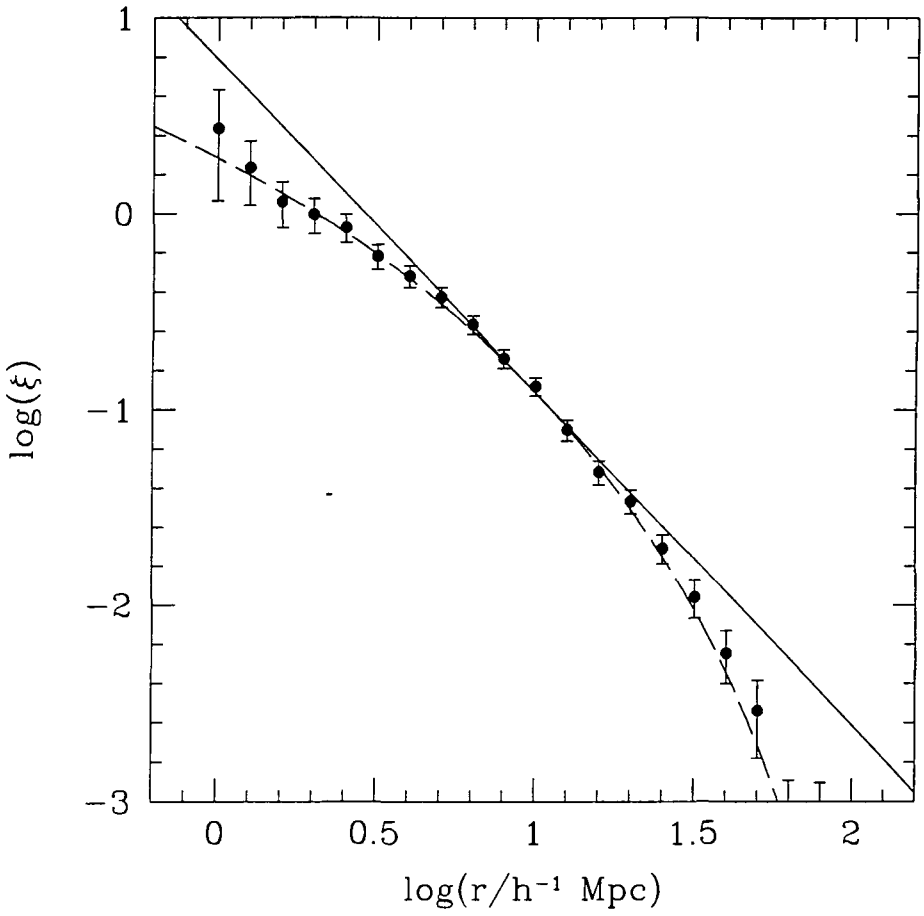

Figure 5.1: The points show the correlation function of the mass in redshift space from the *Hubble Volume* Simulation. The mass particles have redshifts in the range $0< z <2.4$ and have not had the $N(z)$ distribution of the QSOs imprinted on them. Each 75×5 degree slice of the simulation contains 125,000 particles. The dashed line shows the real space correlation function found from the Fourier Transform of the real space input power spectrum. This is scaled to match the amplitude of the mass correlation function in redshift space, assuming that the effects of redshift space distortions and measuring the mass correlation function over a range of different redshifts just alter the amplitude of the clustering. On small scales there should be differences between the two correlation functions due to redshift space distortions and non-linear clustering, which are accounted for in $\xi(r)$ from the mass but not in the $\xi(r)$ inferred from the input power spectrum. The solid line shows a power law of the form $\xi = (r/2.7h^{-1}\mathrm{Mpc})^{-1.7}$, which matches the amplitude of the mass correlation function at $10h^{-1}$Mpc. The value of $r_0 = 2.7h^{-1}$Mpc agrees with the amplitude of clustering of the dark matter at $z = 1.4$ (Figure 5.2). The dark matter correlation function does not have a power law shape.

than from the initial density field. The bias probability is given by

$$P(\nu) = \begin{cases} \exp(\alpha\nu + \beta\nu^{3/2}) & \text{if } \nu \geq 0 \\ \exp(\alpha\nu) & \text{otherwise,} \end{cases} \tag{5.1}$$

where ν is the number of standard deviations of the cell density away from the mean cell density. The parameters α and β are discussed below.

The biasing prescription is as follows:

1) Select one of the three 75×5 degree slices.

2) Split the survey into bins of 0.2 in redshift. Bin the redshift space density field in each redshift bin onto a grid of cell size $20h^{-1}$ Mpc.

3) Select one of the redshift bins and calculate the mean cell density.

4) Select a value of β. We keep α fixed as β is the parameter which most affects the amplitude of the clustering. The larger the value of β, the higher the probability of selecting particles from the densest regions. In order to recover a fixed clustering amplitude, β has to increase as the redshift increases (as the dark matter clustering amplitude decreases). We find that β varies between -1 and 0. We fix α to be 0.15, which is the value given in Cole et al. (1998) such that the real space clustering from a ΛCDM simulation at $z = 0$, with similar parameters to those of the *Hubble Volume*, matches the APM correlation function of Baugh (1996).

5) Calculate the probability of selecting a mass particle to be a biased particle using equation 5.1. Generate a random number between 0 and 1. If this number is less than $P(\nu)$ the particle is a biased particle.

6) Randomly select up to 30,000 biased particles and measure the correlation function.

7) Measure the value of r_0 from the correlation function over the range $5 < r < 30h^{-1}$Mpc, assuming γ=-1.7. If the value matches $6.5h^{-1}$Mpc to within 5%, then go back to 3) and repeat for the next redshift bin. If r_0 is too low then increase β, if r_0 is too high, decrease β. Go back to 4).

This is then repeated for each 75×5 degree slice.

The dark matter correlation function does not have a power law form (Figure 5.1) so we are only able to bias the particles to match approximately the clustering in the 2dF QSO Survey unless we apply a biasing scheme that changed the shape of the dark matter correlation function as well as the amplitude. However, this should be adequate for testing the estimators. In any case, the correlation function from the final 2dF QSO catalogue may have a slightly different shape to the correlation function from the Jan 2000 catalogue.

The results of applying this biasing scheme can be seen in Figure 5.2. We fix γ=-1.7 and find the average value of r_0 in each redshift bin from the three slices. To find the error for each redshift bin, we sum together in quadrature the uncertainty in measuring r_0, over the range $5 < r < 30h^{-1}$Mpc, from each strip. The squares show the clustering amplitude of the dark matter particles and the circles show the clustering amplitude of the biased particles (mock QSOs). The dashed line shows the linear theory prediction for the dark matter clustering as a function of redshift. As designed, the clustering of the mock QSOs is constant with redshift.

The biasing scheme has little effect on the size of the velocity dispersions measured from the biased particles. Figure 5.3 compares the measurement of the rms pairwise velocities from the dark matter particles in the *Hubble Volume* and the biased particles as a function of redshift. The average value of the rms pairwise velocity dispersion at the average redshift of the 2dF QSO Survey is $< w_z^2 >= 400$ km s^{-1}.

5.2.4 Matching the Radial Selection Function

For each strip of the *Hubble Volume* simulation, we have created a catalogue of biased particles that have similar clustering to the 2dF QSO Survey. However, each catalogue contains more than 12,500 QSOs, the number of QSOs expected to be observed in each strip, and the radial selection function differs from that measured from the 9,000 QSOs in the Jan 2000 catalogue. To match this, there are several steps we must take.

We fit a polynomial to the $N(z)$ distribution of the QSOs with redshifts in the range $0.3 < z < 2.4$. We truncate the QSO catalogue at z=0.3 and 2.4 as the completeness of QSOs within this redshift range is better than ~90%. At higher and lower redshifts, the completeness falls rapidly. We renormalise the polynomial such that $\int_{0.3}^{2.4} P(z)dz = 1$ to generate the probability distribution function, $P(z)dz$, of finding a QSO with a redshift in the range z to $z + dz$. For each particle in the biased catalogue, we calculate its value of z from its radial distance from the observer, assuming Ω_m=0.3, Ω_Λ=0.7 and generate a random number between 0 and 1. If the random number is less than $P(z) \times 12,500 \times (2.4 - 0.3)/N_{\rm bias}$ where 12,500 is the approximate number of QSOs that will be observed in each strip and $N_{\rm bias}$ is the number of particles in the biased catalogue of each strip, then the particle is a mock QSO. This generates a catalogue of mock QSOs that has the correct radial selection function and contains approximately the correct number of QSOs. Figure 5.4(a) shows the $N(z)$ distribution of QSOs (histogram) scaled to contain 12,500 QSOs. The dashed line shows the $N(z)$ distribution for mock QSOs on

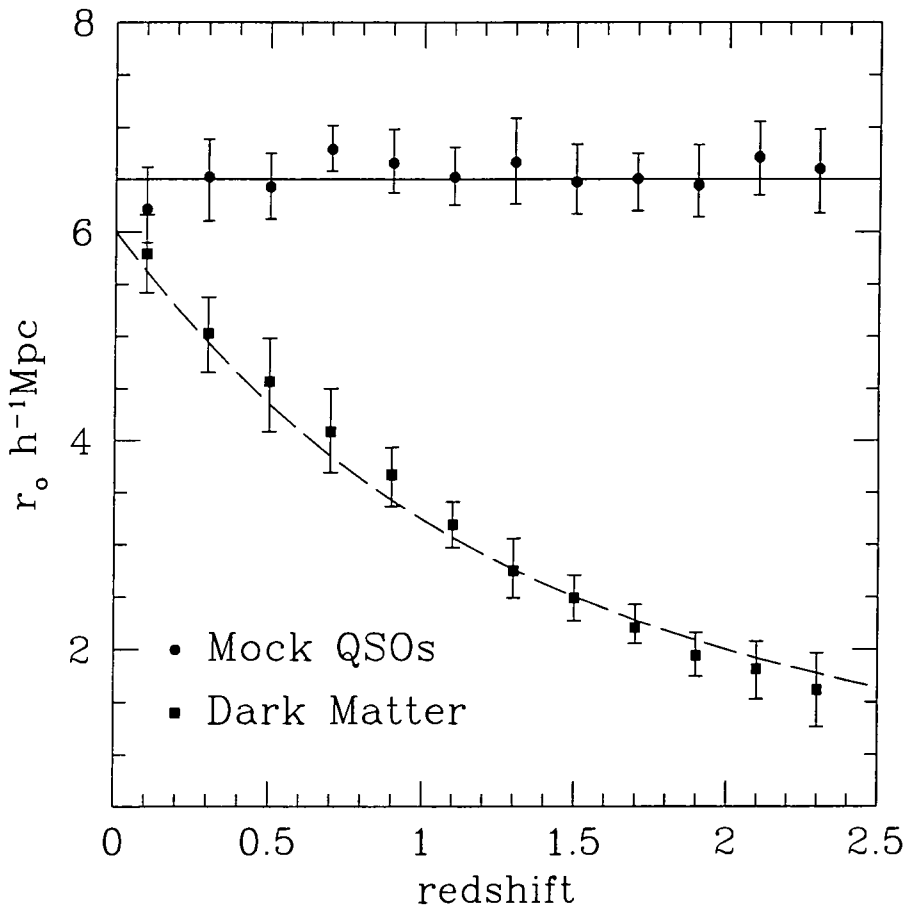

Figure 5.2: The average value of the comoving clustering amplitude, r_0, from the three strips, within each redshift bin, is shown for the biased particles (circles) and for the dark matter particles (squares). The dashed line shows the linear theory prediction for the dark matter clustering as a function of redshift. We assume the correlation function is a power law over the range $5 < r < 30h^{-1}$Mpc with γ=1.7. The errors show the uncertainty in measuring r_0 from each strip over the range $5 < r < 30h^{-1}$Mpc, from the three strips.

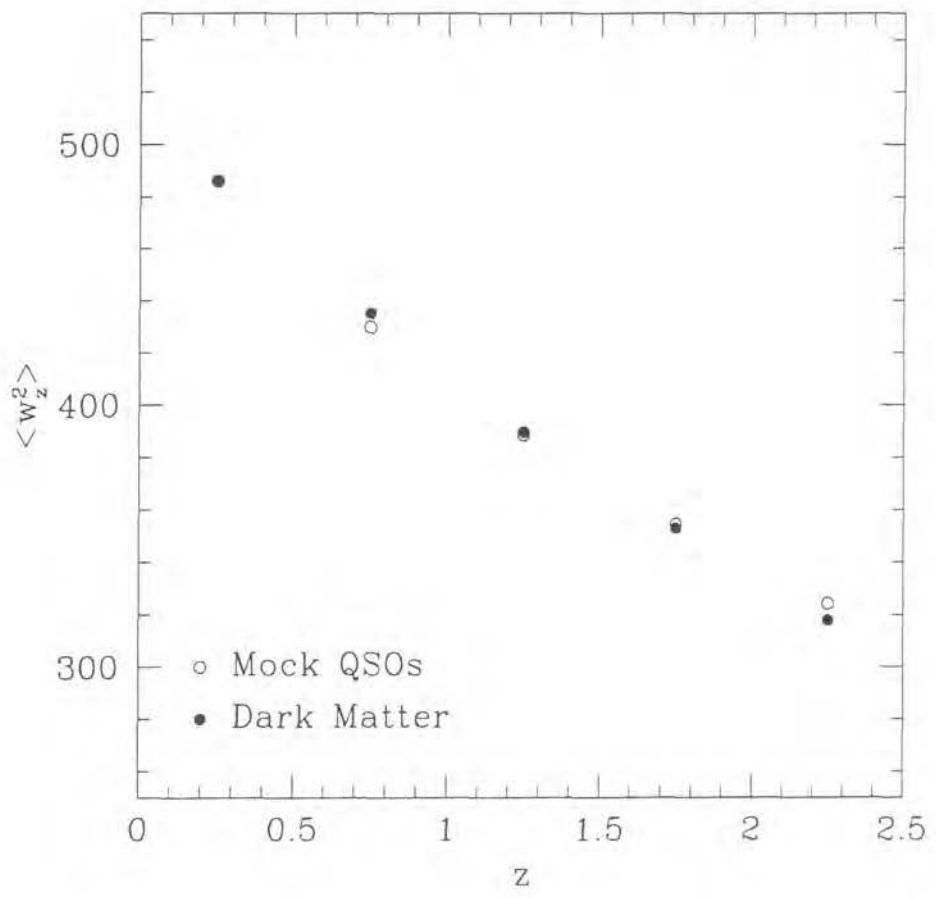

Figure 5.3: The points show the rms pairwise velocity dispersions of the dark matter particles (open circles) in the *Hubble Volume* simulation and the biased particles (closed circles) as a function of redshift. The two sets of particles have very similar values of the rms pairwise velocity dispersion. At the average redshift of the 2dF QSO Survey, the particles have a value of $< w_z^2 > \sim 400$ km s^{-1}.

one 75×5 degree strip.

Figure 5.4(b) shows the number density of one of the mock catalogues with redshifts in the range $0.3 < z < 2.4$ as a function of distance from the observer in bins of $\Delta r = 10h^{-1}$Mpc. The number density varies little as a function of r, the difference being less than a factor of three over the whole redshift range (the number density of the QSOs in the January 2000 catalogue is shown in Figure 6.1). As a comparison, the number density of galaxies in the Durham/UKST survey decreases by a factor of ∼1,000 over a far smaller range of scales (see Figure 2.3 in Chapter 2). This means that all QSOs can be given equal weight when calculating the correlation function or power spectrum. The effect of introducing a weighting scheme when measuring the correlation function or power spectrum, such as that of Feldman, Kaiser & Peacock (1994) (discussed in Chapter 3 in Section 3.2.1) is negligible.

In order to calculate the statistics considered here, we must also generate a catalogue of random points that are unclustered, have the same angular and radial selection function as the survey but have a far higher number density. To create this, we generate a large number (typically 10 million so that the final random catalogue contains 1 million particles) of points in a box that is $3{,}000h^{-1}$Mpc on a side, the same size as the *Hubble Volume* simulation. We select one vertex to be the observer and reject particles that do not lie in the 75×15 degree lightcone that is oriented along the maximal diagonal. We also reject particles that have redshifts out of the range $0.3 < z < 2.4$. This matches the angular selection function of the 2dF QSO Survey.

To match the radial selection function, we use the same method as described above. The redshift of each particle is calculated, again assuming $\Omega_{\rm m}$=0.3, Ω_Λ=0.7 to convert from r to z. We use the probability distribution function found from the polynomial fit to the $N(z)$ distribution of QSOs to calculate the probability of the particle remaining in the random catalogue. The only difference is that the term $12,500/N_{\rm bias}$ is removed in the calculation of the probability of selecting a particle as we do not sparse sample the random catalogue.

5.3 The Two-Point Correlation Function $\xi(r)$

5.3.1 Estimation of $\xi(r)$

One of the most widely used measurements of clustering is the two-point correlation function (for the correlation functions of surveys considered in this Thesis see Loveday

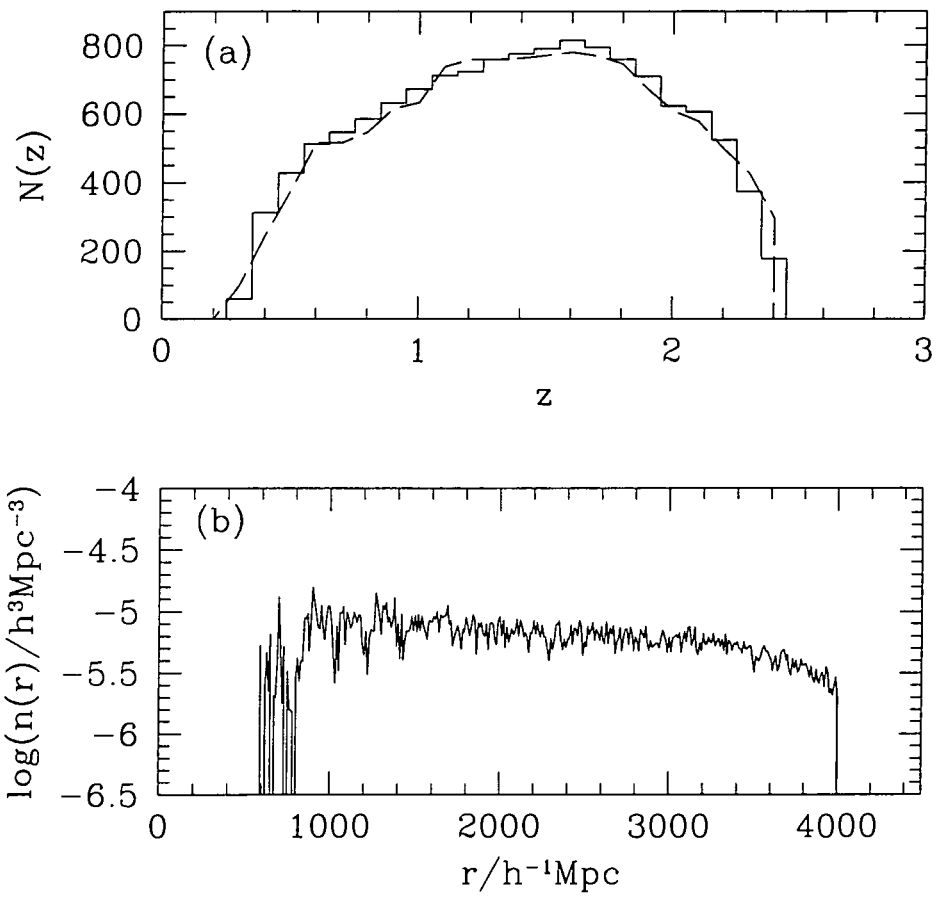

Figure 5.4: Panel (a) shows the number of QSOs in bins of 0.1 in redshift, $N(z)$, for QSOs with redshifts in the range $0.3 < z < 2.4$ (histogram). This is calculated from the first 9,000 QSOs to be observed and scaled to show the predicted $N(z)$ distribution for 12,500 QSOs on each strip. The dashed line shows the $N(z)$ distribution of one of the mock catalogues. QSOs are binned into the same 0.1 bins of redshift and have redshifts over the same range, $0.3 < z < 2.4$. Panel (b) shows the radial number density of biased particles in one of the mock catalogues in bins of $\Delta r = 10h^{-1}$Mpc. The number density is almost constant out to very large scales, allowing a simple weighting scheme to be adopted when measuring clustering statistics.

et al. 1992, Croom & Shanks 1996, Baugh 1996, Ratcliffe et al. 1998b). The two-point correlation function is easy to calculate, it can be calculated from small surveys and it is perhaps easier to interpret than the power spectrum. A high value of the correlation function means strong clustering on that particular scale, as long as the bins are not correlated.

Following Peebles (1980), the two-point QSO-QSO correlation function, $\xi(r)$, is defined as the excess probability of finding two objects, separated by a distance r in two volume elements $\delta V_i, \delta V_j$

$$\delta P_{i,j}(r) = \bar{n}\bar{n}[1 + \xi(r)]\delta V_i \delta V_j \tag{5.2}$$

where $\bar{n}$ is the mean density of QSOs in the survey. The value of ξ determines the properties of the clustering. If $\xi > 0$ then the distribution is clustered, if $\xi < 0$ then the distribution is anti correlated, i.e. there are fewer pairs at a given separation than would be expected in a random distribution, which is characterised by $\xi = 0$.

In order to measure the correlation function, equation 5.2 must be changed into a more practical form. The mean number of QSOs expected in a volume $\Delta V_{i,j}$ at redshift z_j at a separation r away from a randomly chosen QSO i in a flux limited sample is given by

$$< N_{i,j}(r) >= n(z_j)(1 + \xi(r))\Delta V_{i,j} \tag{5.3}$$

where $n(z_j)$ is the mean number density at the redshift of QSO j. The expected pair count, found by summing over all volumes $\Delta V_{i,j}$ at a separation r from i that lie within the survey is given by

$$< DD(r) >= \left[\sum_i \sum_j n(z_j)\Delta V_{i,j}\right](1 + \xi(r)). \tag{5.4}$$

If particles are randomly distributed around the QSOs, such that they have the same radial density profile but no clustering, the number of QSOs and random particles expected in any bin of size r, $DR(r)$, is given by the expression $\left[\sum_i \sum_j n(z_j)\Delta V_{i,j}\right]$. Therefore the two-point correlation function, $\xi(r)$ can be *estimated* as (Peebles 1980)

$$\xi(r) = \frac{DD(r)\bar{n}_R}{DR(r)\bar{n}_D} - 1, \tag{5.5}$$

where $DD(r)$ is the number of QSO-QSO pairs at a given separation, r, and $DR(r)$ is the number of QSO-random pairs. The $DD(r)$ counts are found by auto correlating the data catalogue with itself and the $DR(r)$ counts are similarly found by cross correlating the data catalogue with a catalogue of random points. The random catalogue has the

same angular and radial selection function as the data catalogue but without any intrinsic clustering. The construction of such a random catalogue was discussed in Section 5.2.4. $\bar{n}_D$ and $\bar{n}_R$ normalise the counts if the number of data points is different to the number of random points. At least a factor of 10 more random points than data points are used and typically 300,000 random points (a factor of 24 more) are used to define the geometry of each slice of the Survey. This is sufficient to describe the survey geometry whilst keeping the time to compute $\xi(r)$ reasonably short.

The correlation function estimate, equation 5.5, may be affected by the integral constraint. This constraint arises because both the pair counts and the mean density of QSOs in the Universe have to be estimated from the survey itself. The integral constraint causes the correlation function to be biased low The mean density of QSOs in the Universe cannot be properly measured from the survey as it is convolved with the large scale structure within the survey. The volume of the 2dF QSO Survey is large but even so, the survey may not be truly representative of the whole Universe as the area of sky covered by the survey is still quite small. The mean density of the survey may be sensitive to whether structures are fully sampled by the survey. The simple estimator of the correlation function (equation 5.5) was therefore improved upon by Hamilton (1993), who removed the need to know the correct mean density by using the estimator

$$\xi(r) = \frac{DD(r)RR(r)}{DR(r)^2} - 1, \tag{5.6}$$

and by Landy & Szalay (1993) who also improved the simple estimator by expanding out the expression $\xi = (D - R)^2/R^2$ to define

$$\xi(r) = \frac{DD(r) - 2DR(r) + RR(r)}{RR(r)}. \tag{5.7}$$

In this case, the DR and RR counts still have to be correctly normalised to the DD counts in the Landy and Szalay estimator in a similar way as for the simple estimator, however the bias introduced from this method is shown to be only a second order effect in Landy & Szalay (1993).

The upper panel in Figure 5.5 shows the estimates of the redshift space correlation function from the mock catalogues, found using the three methods described here. The solid line in Figure 5.5 shows the latest measurement of the correlation function from the 2dF QSO survey. The simulated correlation function has approximately the same amplitude and shape as the current sample of ~9,000 QSOs. Unless the first QSOs to be observed have a significantly different clustering pattern to the rest of the QSOs that will

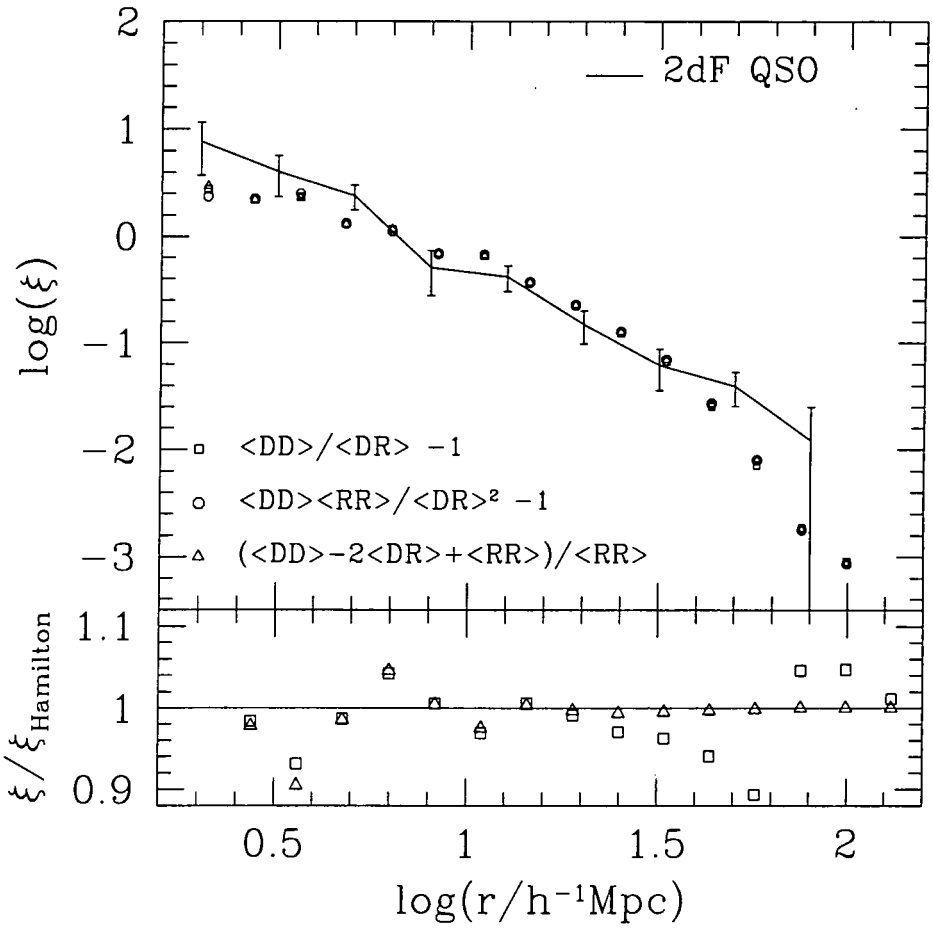

Figure 5.5: The upper panel shows the average correlation function measured from the three mock catalogues which have the same angular and radial selection function, contain the same number of mock QSOs and have approximately the same clustering as that expected in the final 2dF QSO Survey. The correlation functions are estimated using the $\xi(r)$ estimator of Peebles (1980, squares), Hamilton (1993, circles) and Landy & Szalay (1993, triangles). The solid line shows the latest correlation function from the 2dF QSO survey, measured from ~9,000 QSOs (from Scott Croom). All the correlation functions are in *redshift* space. The lower plot shows the estimation of the correlation function using the estimator of Peebles (1980) (triangles) and Landy & Szalay (1993) (open circles), normalised to the correlation function measured using the estimator of Hamilton (1993). The measurements of the correlation function using the estimators of Hamilton and Landy and Szalay are in excellent agreement. The simple estimator of Peebles matches the estimator of Hamilton to ~10%.

be observed in the 2dF QSO survey, the mock catalogues should be realistic in the sense that the two-point correlation functions are similar.

The lower panel in Figure 5.5 shows the correlation function measured using the estimator of Peebles (1980) and Landy & Szalay (1993) divided by the correlation function measured using the estimator of Hamilton (1993). There is very good agreement between the correlation functions from the estimators of Hamilton and Landy and Szalay and even the simple estimator of Peebles agrees to $\sim$10%. In order to reduce the chance of systematic errors creeping into the estimation of the two-point correlation function of the mock catalogues we use the estimator of Hamilton.

We have tested the program that calculates the correlation function by measuring the correlation function of the mass and we are able to reproduce the shape of the correlation function inferred from the input power spectrum. The input power spectrum is calculated at $z = 0$ in real space rather than at $z = 1.4$ in redshift space. On linear scales, the difference between a correlation function at $z = 0$ and $z = 1.4$ is just an amplitude shift. The difference between a real space and redshift space correlation function is also just an amplitude shift on scales where the small scale peculiar velocities have a negligible effect on the clustering. This is the case on scales larger than $\gtrsim 5h^{-1}$Mpc, found by comparing the shape of the redshift space mass correlation function measured from the *Hubble Volume* simulation to the real space correlation function inferred from the Fourier Transform of the input power spectrum (Figure 5.1).

5.3.2 Errors

There are many different ways of estimating the errors on $\xi(r)$ analytically. One method of checking which analytic approach is the most appropriate is to use mock catalogues that have a similar clustering pattern as the data. The error is then found by measuring the dispersion over a large number of these mock catalogues. This was the approach taken in Chapter 3. Unfortunately, there are only three mock catalogues available as large amounts of computing resources are required to simulate the volume of the 2dF QSO Survey. The errors found from the dispersion over the mock catalogues should give the overall shape of the error but they show large fluctuations from scale to scale. We therefore see if other error estimators predict errors that are similar to the mock catalogue errors but which are more stable from scale to scale.

We consider three possible analytic expressions for the errors on the two-point correlation function. These are Poisson errors, bootstrap errors and, what we term, constant

N_q errors. Poisson errors are defined by

$$\Delta\xi(r) = (1+\xi)\sqrt{\frac{2}{DD(r)}}. \tag{5.8}$$

Note, the factor of two is included as we count each QSO-QSO pair twice. On small scales, where the number of pairs in each bin is also small and most pairs are independent, Poisson errors are thought to provide a good estimation of the errors. On large scales, the pairs are not independent and Poisson errors can underestimate the true error.

Bootstrap errors (Ling, Frenk & Barrow 1986) are found by constructing random samples of the catalogue that have the same number of data points as the real catalogue (i.e. some data points are included more than once, some not at all), and then calculating the correlation function from these samples. The error is then the dispersion found in the DD counts from a large number of these samples, scaled to the correlation function. 50 samples were used for each of the three mock catalogues, giving a total of 150 samples.

A final type of error is the constant N_q errors. On large scales, Shanks & Boyle (1994) and Croom & Shanks (1996) found that Poisson errors underestimated the error found from field-to-field variations. Instead, they found that a better agreement was found if the factor $DD(r)/2$ in equation 5.8 was replaced by the total number of objects in the survey, which here corresponds 25,000. On small scales, these errors drastically underestimate the error found from all the other methods as there are far fewer than 25,000 pairs in each DD bin. On large scales, the pairs in each DD bin are not independent and therefore the errors are underestimated if $DD(r)/2$ is used in the error estimation.

All the errors are summarized in Figure 5.6. The mock catalogue errors are shown by the solid line. The fluctuation in the error from scale to scale can be clearly seen. The long dashed line shows the Poisson errors and the short dashed line shows the Bootstrap errors. The dotted line shows the constant N_q errors which clearly underestimate the errors on small scales.

On small scales, Poisson errors seem to match the errors from the mock catalogues most closely. Bootstrap errors seem to over estimate the error and constant N_q errors drastically under estimate the error for reasons mentioned above. Bootstrap errors are approximately $\sqrt{3}$ larger than the Poisson errors, consistent with the findings of Mo, Jing & Börner (1992).

On large scales, the constant N_q errors seem to overestimate the errors. However, the mock catalogues are not independent and on large scales may be underestimating the error. We therefore take the conservative view that the constant N_q errors give a

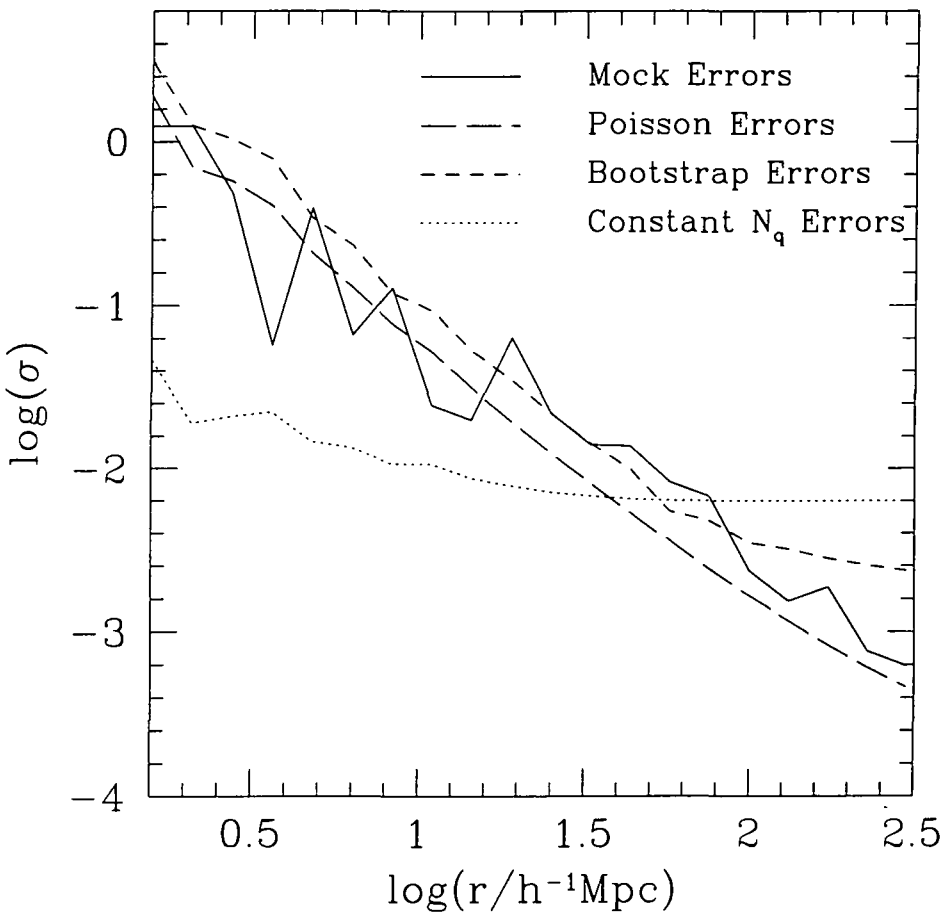

Figure 5.6: The error on the correlation function for different error estimates: dispersion over the 3 available mock catalogues (solid line), Poisson errors (long dashed line), errors found by bootstrap resampling 150 times (short dashed line) and constant N_q errors (dotted line).

reasonable estimation of the error.

However, on intermediate scales, $20 \lesssim r \lesssim 60h^{-1}$Mpc, Poisson and N_q errors seem to underestimate the errors for the mock catalogues. Bootstrap errors seem to give the closest match to the mock catalogue errors but these are time consuming to calculate. Previously, the rule of thumb has been to assume Poisson errors up to the scale where $DD(r)/2 \sim N_q$ and on scales larger than this, constant N_q errors were found to provide the best error estimator (Croom & Shanks 1996) and we will adopt this method here. However, more independent mock catalogues are really required to allow us to determine which technique matches the errors from mock catalogues most closely. Ideally, we would like enough mock catalogues that the errors found from the dispersion over the mock catalogues converged and different estimations of the errors would not be required but simulations large enough to cover the volume of the 2dF QSO Survey take many hours of computing time to run. An alternative possibility may be to use the Zeldovich approximation as a way to obtain errors in the linear and mildly non-linear regime as a large number of these simulations can be run in a relatively short period of time.

5.3.3 Predicted Correlation Function

Figure 5.7 shows the redshift space correlation function of the mock 2dF QSO survey using the Hamilton estimator, with Poisson errors on scales less than $40h^{-1}$Mpc and constant N_q errors on larger scales. The upper plot shows the correlation function plotted in log-log space whereas the lower plot shows the correlation function plotted in linear-log space. The upper plot shows that the correlation function will be measured out to large, $\sim 100h^{-1}$Mpc scales, before the signal starts to oscillate around zero, although the actual scale will depend on the clustering in the 2dF QSO Survey. The single open point indicates where the correlation function is negative and its absolute value is shown. With a bin size of $\Delta\log(r/h^{-1}\mathrm{Mpc})$=0.12, the errors in our model are around 20% over the range $3 \lesssim r \lesssim 50h^{-1}$Mpc and are around 10% over the range $10 \lesssim r \lesssim 30h^{-1}$Mpc. The solid line in both plots has the form $\xi(r) = (r/6.5h^{-1}\mathrm{Mpc})^{-1.7}$ as a reference, although the correlation function is more consistent with a double power-law, as seen earlier in Figure 5.1 and 5.5.

The lower plot shows the correlation function plotted out to large scales and shows that the correlation function can be measured out to $1{,}000h^{-1}$Mpc. This was one of the key aims of the 2dF QSO Survey. The errors on these scales are very small, constant N_q errors have a value of 0.0063 when ξ=0. This is a factor of 5 improvement on previous

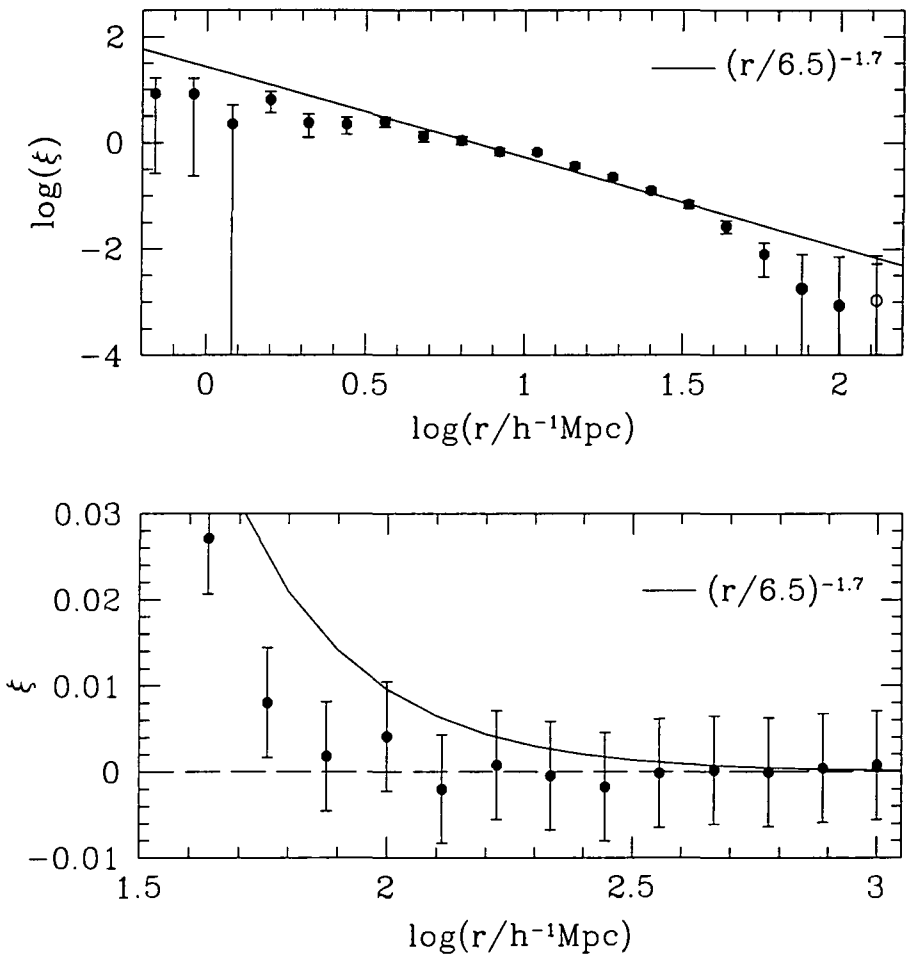

Figure 5.7: The correlation function of the mock 2dF QSO survey measured in redshift space. The upper plot shows the correlation function plotted in log-log space out to $\sim 100h^{-1}$Mpc whereas the lower plot shows the correlation function plotted in linear-log space on large scales. The errors are Poisson errors on scales less than $40h^{-1}$Mpc and constant N_q errors on scales larger than this. The correlation function of the mock catalogues becomes negative only on scales greater than $100h^{-1}$Mpc, indicated by the open symbol in the upper panel where we have plotted $\log|\xi(r)|$. The lower panel shows just how accurately the correlation function will be measured on large scales. The constant N_q errors have a value of 0.0063 when ξ=0.

surveys (Croom & Shanks 1996).

5.3.4 The Evolution of Clustering

We also wish to know how QSO clustering evolves as a function of redshift. This can be determined by splitting the QSO survey into a number of redshift bins and measuring the clustering within each bin. This allows the QSO-mass bias as a function of redshift to be measured relative to the QSO-mass bias at $z = 0$ as a function of cosmology. By measuring the evolution of QSO clustering, we may be able to rule out various models for QSO bias such as those of Matarrese et al. (1997), Colín et al. (1999) and Percival & Miller (1999). To determine the actual QSO-mass bias, further information, such as the clustering amplitude of the dark matter at $z = 0$, is required. One way of determining this is described in more detail in Chapter 7.

Our favoured statistic for measuring the clustering evolution is to use the volume averaged correlation function. The volume averaged correlation function is less noisy than the correlation function measured over a range of scales. It is estimated by finding all the DD, DR and RR pairs that have a separation of $r < r_{\rm max}$. The Hamilton estimator is then used to find $\bar{\xi}$ in the same way as for the two-point correlation function.

The mock catalogues have been constructed under the assumption that QSO clustering is approximately constant with redshift. This behaviour was found by Croom & Shanks (1996) in existing surveys, assuming $\Omega_{\rm m}$=1, and early results from the 2dF QSO survey suggest that this is approximately the case for the clustering in the 2dF QSO Survey (Croom et al. 2000). We also look at the evolution of clustering of dark matter particles that have the same angular and radial selection function and number density that we predict for the 2dF QSO Survey but that are unbiased so their clustering amplitude decreases with redshift in a predictable manner. We use our mock 2dF QSO catalogues here to test how accurately the evolution of QSO clustering will be measured as a function of redshift in the $\Omega_{\rm m}$=0.3, Ω_Λ=0.7 cosmology. We split the mock and dark matter catalogues into 4 bins and measure the clustering and the errors within each redshift bin. The bins are $0.3< z <1.0$, $1.0< z <1.4$, $1.4< z <1.8$ and $1.8< z <2.4$, and are chosen so that each bin contains approximately the same number of objects.

Figure 5.8 shows the volume averaged correlation functions of the biased, mock 2dF QSO catalogues (circles) and the same for the dark matter particles (squares) from each redshift bin. The solid line shows the volume averaged correlation function, $\bar{\xi} = (3/r_{\rm max}^3) \int_0^{r_{\rm max}} \xi(r) r^2 dr$, with $r_{\rm max} = 20h^{-1}$Mpc, assuming $\xi(r) = (r/6.5h^{-1}{\rm Mpc})^{-1.7}$.

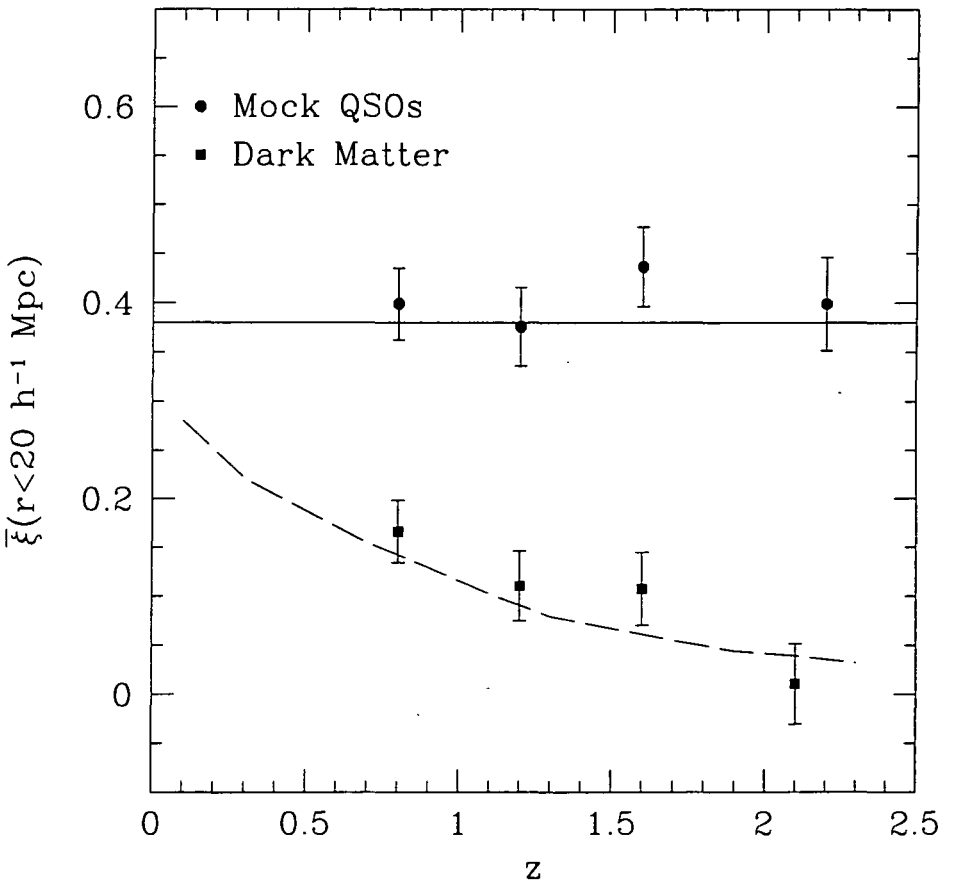

Figure 5.8: The volume averaged correlation function with $r_{\max} < 20h^{-1}$Mpc of the mock 2dF QSO survey (circles) and the dark matter particles (squares), both of which have the same angular and radial selection function as expected in the completed 2dF QSO Survey and are sparse sampled to contain $\sim 12,500$ particles on each 75×5 degree slice. The redshift bins are $0.3 < z < 1.0$, $1.0 < z < 1.4$, $1.4 < z < 1.8$ and $1.8 < z < 2.4$, chosen so that each bin contains approximately the same number of QSOs. Note that $\bar{\xi}$ is plotted linearly. The solid line shows the value of $\bar{\xi}(r < 20h^{-1}\text{Mpc})$ expected from $\xi(r) = (r/6.5h^{-1}\text{Mpc})^{-1.7}$ and the dashed line shows the value of $\bar{\xi}(r < 20h^{-1}\text{Mpc})$ expected from $\xi(r) = (r/\bar{r}_0)^{-1.7}$ with $r_0(z)$ for the dark matter taken from Figure 5.2.

The dashed line shows the volume averaged correlation function, found in the way described above, but with r_0 taken from Figure 5.2, calculated in bins of 0.2 in redshift. As expected, the clustering is approximately constant for the biased particles and decreases for the dark matter particles as redshift increases. The errors shown are Poisson errors found from the DD counts as $r_{\rm max}$ is 20 h^{-1}Mpc, Section 5.3.2.

If the 2dF QSO survey is split into 4 bins, then the errors on the measurement of the volume averaged correlation function are ~10% for the mock catalogues used here. This should allow the relative QSO-mass bias as a function of redshift and cosmology to be determined. Even if the clustering amplitude of QSOs decreases like the dark matter, we should still be able to measure the clustering in the highest redshift bin. Early results from the 2dF QSO Survey (Croom et al. 2000) and previous results on QSO clustering (Croom & Shanks 1996) suggest that it is unlikely that the clustering of QSOs will evolve in this way.

5.4 The Power Spectrum

5.4.1 Introduction to the power spectrum

As discussed in Chapter 3, the power spectrum is the Fourier transform of the correlation function. However, when the correlation function or power spectrum is measured from a finite and noisy data set, the two are not an exact Fourier conjugate pair. To extract as much information from a Survey as possible, it is worth considering both the correlation function and the power spectrum. The main advantage of the power spectrum is that uncertainties in the mean QSO density only effects the power spectrum on the very largest scales (Cole, Fisher & Weinberg 1995), whereas uncertainties in the mean density of QSOs affects the correlation function on all scales, although Hamilton (1993) and Landy & Szalay (1993) have improved the estimation of the correlation function. The power spectrum is also the quantity predicted directly by theory, allowing comparisons between models of Large Scale Structure and power spectra measured from galaxy or QSO redshift surveys to be made easily. An additional advantage is that the power spectrum is also quick to compute using Fast Fourier Transforms (FFT).

The power spectrum of optically selected QSOs has never been measured before. Existing QSO surveys have a large volume, due to the great depths to which QSOs can be detected. However, they cover only a small area of sky and contain a small number of QSOs. This restricts the range of scales over which the power spectrum could be reliably

measured due to the effects of the survey window function on large scales and shot noise on small scales, if a FFT is used to measure the power spectrum. The greater solid angle and the vastly increased number of QSOs in the 2dF QSO survey make it a better survey for power spectrum analysis. In Chapter 6, we use the techniques developed in this Chapter to measure the power spectrum from the first 9,000 QSOs to be observed in the 2dF QSO Survey.

In this Section, we first of all examine how the geometry of the survey affects the power spectrum and check that measuring the power spectrum from QSOs over a wide range of redshifts does not introduce any systematic effects into the shape of the power spectrum. The method of estimating the power spectrum, either calculating the Fourier transform directly or using a Fast Fourier Transform (FFT) is tested and finally we compare different error estimators. We summarize this section by asking whether the turnover expected in CDM power spectra can be detected.

5.4.2 The Window Function

As discussed in Chapter 3, the power spectrum measured directly from a survey is a convolution of the true power spectrum with the window function of the survey. Ideally the power spectrum of the window function should be deconvolved from the power spectrum of the QSOs or galaxies. This process is complicated and as yet this has only been attempted for the Las Campanas Redshift Survey (Lin et al. 1996). However, the window function only affects the power spectrum shape on the very largest scales. Rather than attempt to deconvolve the two, we shall estimate the scales on which the geometry of the survey alters the shape of the measured power spectrum, which is the same approach that we took in Chapter 3.

The power spectrum of the window function is estimated by placing $\sim$1 million random points into the survey geometry. These particles have the same angular and radial selection function as the observed QSO sample, see Section 5.2.4. One million particles are sufficient to accurately trace out the geometry. This can be tested as the sum of the window function power spectrum is equal to the volume of the box that the survey is embedded into (see Chapter 3 and Section 5.4.3) divided by the volume of the survey. Once this value has converged to an accuracy of a few percent, typically 5%, sufficient random points have been used. More random points can be used in the power spectrum estimation than in the estimation of the correlation function as the power spectrum is quick to compute using a FFT. The time required to compute a power spectrum depends

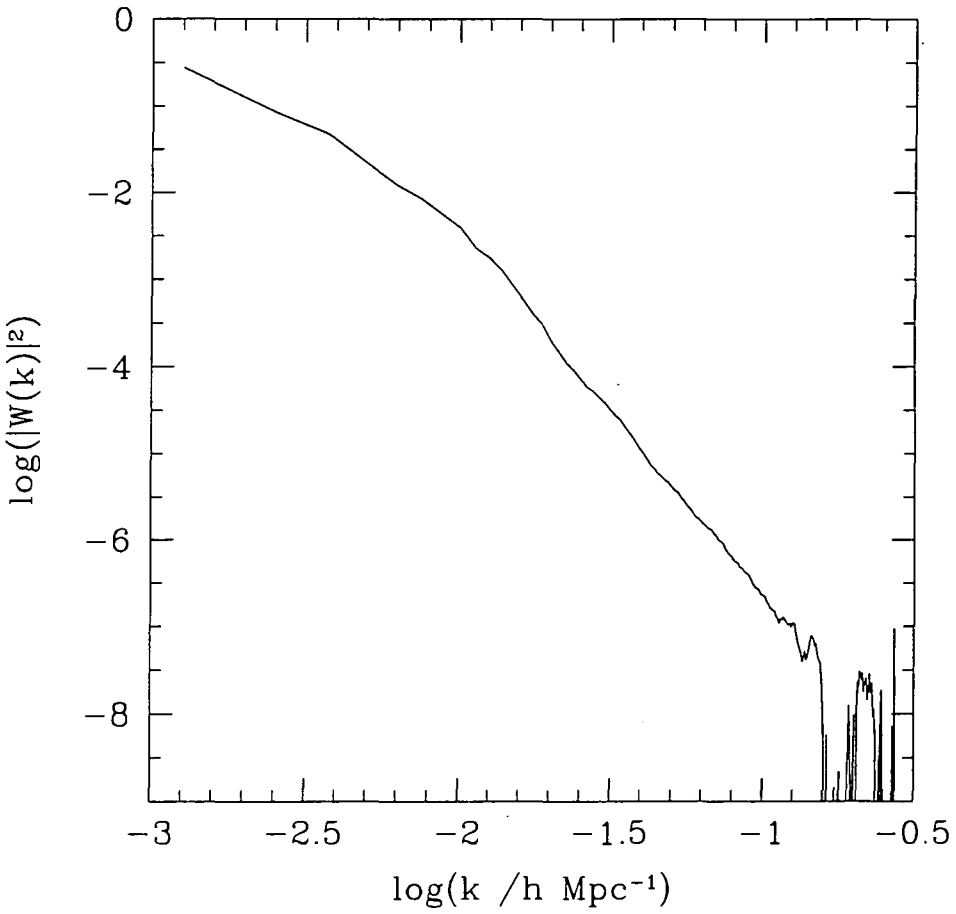

Figure 5.9: The power spectrum of the window function for one slice (5×75 degrees) of the 2dF QSO survey, estimated from a random catalogue that has the same angular and radial selection function as the QSOs but without any clustering. The catalogue contains many more particles than there will be in the final 2dF QSO Survey. For wavenumbers $\log(k/h\mathrm{Mpc}^{-1}) \gtrsim -2$ ($\sim 600h^{-1}\mathrm{Mpc}$) the window function power law is a steep power law $\propto k^{-4}$. The dip at $40h^{-1}\mathrm{Mpc}$ occurs at the Nyquist frequency, discussed later.

more on the size of the Fourier transform grid rather than on the number of data or random points. Figure 5.9 shows the shape of the power spectrum of the window function. The power spectrum is a steep power law proportional to k^{-4} on scales smaller than $\sim 600h^{-1}$Mpc.

To quantify the scales over which the measured power spectrum will be affected by the survey geometry, the power spectrum of the window function is convolved with the power spectrum used to describe the initial conditions of the *Hubble Volume* simulation. This is done as in equation 2.1.6 in Feldman, Kaiser & Peacock (1994) but we approximate the convolution by integrating over k rather than k_x, k_y, k_z. The convolved power spectrum matches the power spectrum measured from the mass, which is fully convolved with the power spectrum of the window function when we compute the FFT, to within the errors, which suggests the approximation is adequate for these purposes.

Figure 5.10 shows the input power spectrum from CMBFAST (solid line) and the effect of convolving this with the power spectrum of the window function (dashed line). The effect of the window function is to change the shape of the power spectrum on scales larger than $400h^{-1}$Mpc. On scales $\lesssim 400h^{-1}$Mpc the input power spectrum is free from any distortions caused by the survey geometry. On scales larger than $400h^{-1}$Mpc the survey geometry changes the shape of the power spectrum but only by a small amount. This effect will be smaller than the measurement errors on these scales (see Section 5.4.6).

5.4.3 Clustering along a Lightcone

The 2dF QSO survey will observe QSOs over a range of redshifts, whereas current wide angle galaxy redshift surveys, even the 2dF Galaxy Redshift Survey, are essentially probing the $z = 0$ space. This means that the QSO clustering will be imprinted along an observer's past lightcone. By considering the input power spectrum, the power spectrum measured from a series of redshift bins and the power spectrum of the mass from the whole simulation lightcone, we show that measuring a power spectrum from a lightcone has little effect on the shape of the power spectrum, at least on scales $60 < r < 400h^{-1}$Mpc, or on the position of the turnover. The amplitude of the power spectrum from the mass in the lightcone is lower than the input linear theory power spectrum calculated at $z = 0$ due to the evolution of clustering in the dark matter.

The solid line in Figure 5.11(a) and (b) is the input power spectrum to the *Hubble Volume* simulation, calculated using CMBFAST (Seljak & Zaldarriaga 1996) at $z = 0$. The five dashed lines in panel (a) are mass power spectra in redshift space from the *Hubble*

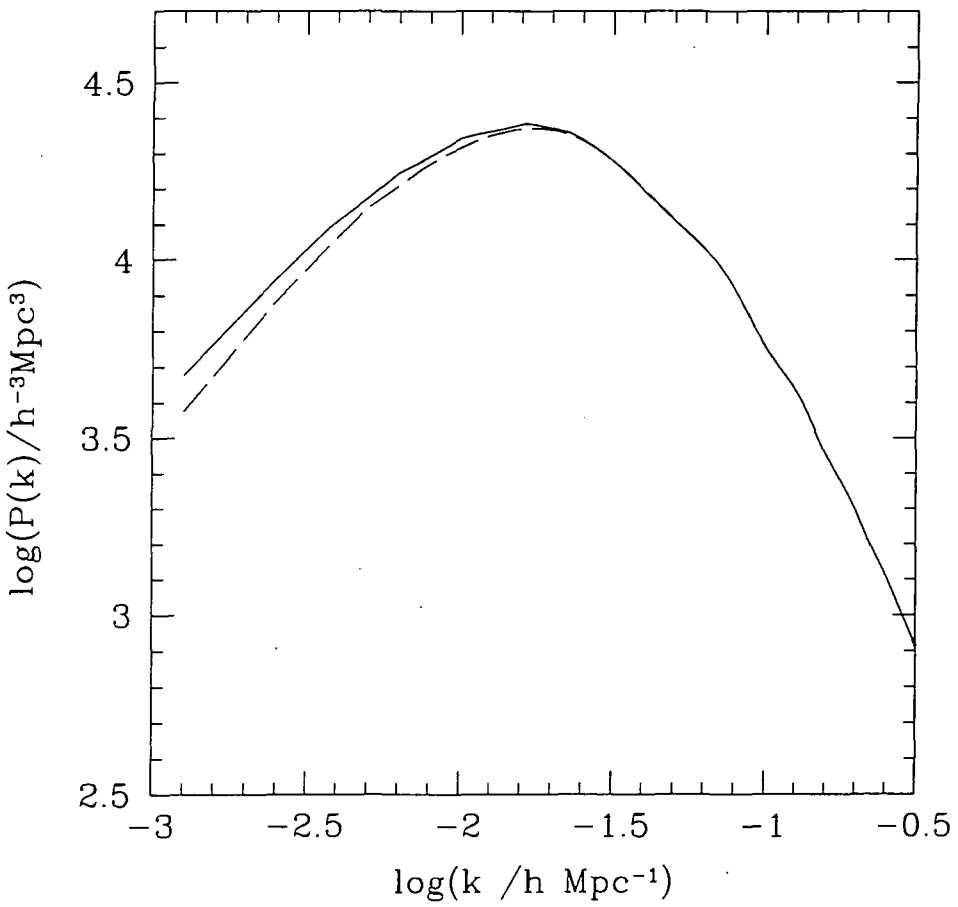

Figure 5.10: The solid line shows the input power spectrum to the simulation and the dashed line shows the input power spectrum convolved with the power spectrum of the window function, shown in Figure 5.9. The method is described in the text. The window function causes deviations away from the true power spectrum on scales greater than around 400 h^{-1}Mpc (log(k/hMpc^{-1})=-1.8).

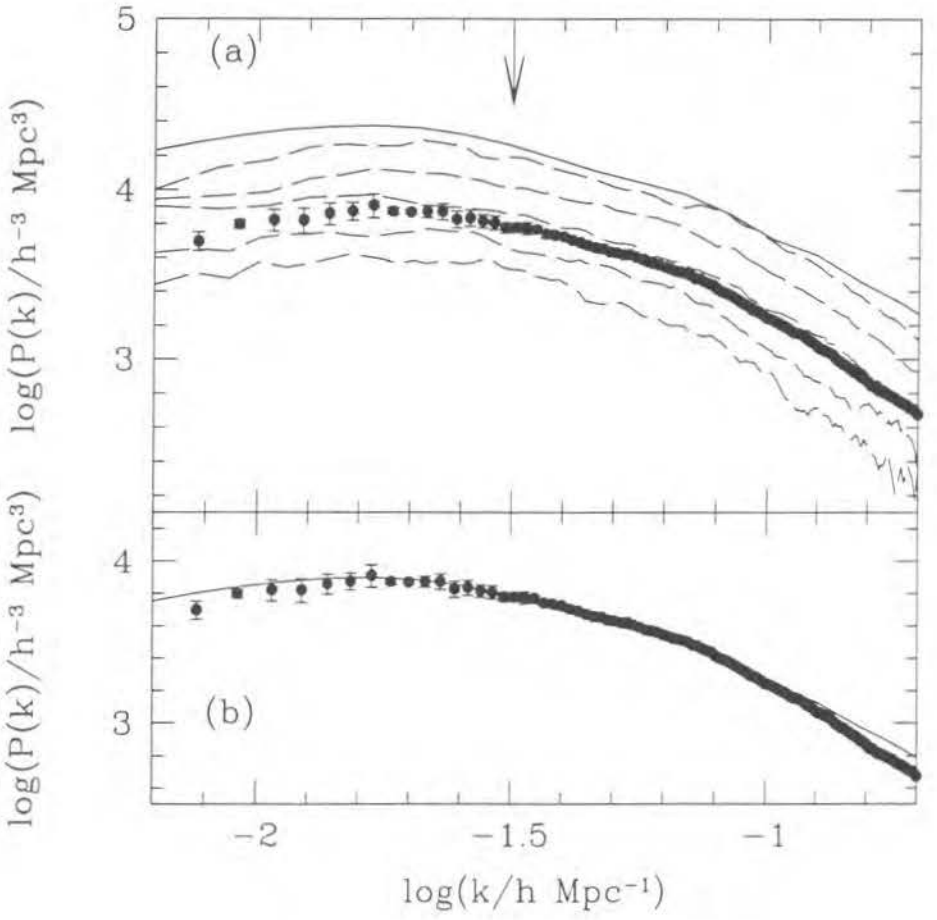

Figure 5.11: The solid line in panels (a) and (b) is the input linear theory real space power spectrum calculated at $z = 0$. The dashed lines in panel (a) are the power spectra of the dark matter in redshift space from bins of 0.5 redshift bins centred on (reading top to bottom) 0.25, 0.75, 1.25, 1.75, 2.25. At this stage, the $N(z)$ distribution of the QSOs has not been imprinted on the dark matter particles. The arrow indicates the scale up to which the power spectrum from the smallest redshift bin is unaffected by the geometry of that bin. The points in panels (a) and (b) show the power spectrum of the dark matter particles in redshift space over the range of redshifts $0 < z < 2.4$. This power spectrum has a similar shape to the power spectra from the $z = 1.25$ and $z = 1.75$ bins and lies between them as the median redshift of particles in the lightcone is $\bar{z} = 1.4$. The input power spectrum in panel (b) has been scaled to match the amplitude of the dark matter power spectrum. The two power spectra have very similar shapes over a wide range of scales and the turnover from a positive to a negative slope occurs on the same scale, $r = 400h^{-1}$Mpc. There are slight discrepancies on the smallest and largest scales, discussed in the text. Measuring the power spectrum from a lightcone is shown here not to alter the shape of the power spectrum, at least on scales $60 < r < 400h^{-1}$Mpc.

spectrum, Figure 5.1, we find that the effects of the small scale peculiar velocities are only important on scales less than $3h^{-1}$Mpc so redshift space distortions caused by peculiar velocities should not affect the shape of the power spectrum on scales of $60h^{-1}$Mpc.

As seen in Chapter 3, the FFT is only reliable down to the Nyquist frequency, defined as

$$k_{\rm Nyquist} = \frac{2\pi}{l_{\rm box}} \frac{\rm N_{grid}}{2}, \tag{5.9}$$

where $\rm N_{grid}$ represents the size of the FFT transform, in our case 256, and $l_{\rm box}$ is the size of the cube that the survey is embedded into in order to carry out the FFT. This can be any size as long as the survey can fit into it. We found in Chapter 3 that the larger the box, the better the large scale modes were sampled. However, the larger the box, the lower the Nyquist frequency, which then limits how well the power spectrum on small scales can be measured. We adopt a value of $4000h^{-1}$Mpc, giving a value of $\log(k_{\rm Nyquist}/h{\rm Mpc}^{-1})$ = -0.7, corresponding to a scale of $30h^{-1}$Mpc. However, scales larger than this can be affected by the size of the FFT. Using the nearest grid point scheme, where particles in the density field are simply assigned to the nearest cell, the power spectrum can only be reliably measured down to around half the Nyquist frequency, checked empirically by Hatton (1999). This corresponds to $\log(k/h{\rm Mpc}^{-1})$ = -1 or a scale of $60h^{-1}$Mpc. In Figure 5.11(b), the mass power spectrum does deviate from the input power spectrum on scales of $\sim 60h^{-1}$Mpc.

By directly calculating the Fourier transform, the small scale power spectrum can be measured more accurately. This is described in Section 5.4.5. However, due to computational limits, the direct Fourier transform is only feasible for $< 50,000$ particles. The direct method can be used to calculate the power spectrum of the real 2dF QSOs (see Chapter 6) or the mock 2dF QSOs as seen in Section 5.4.5.

5.4.4 P(k) from the biased particles

We wish to test how well a linear bias can reconcile the power spectrum of the biased particles with the power spectrum of the mass. As the scheme adopted here biased the dark matter particles to have the same correlation function amplitude on scales of $5 < r < 30h^{-1}$Mpc, a linear bias should exist between the mass and the biased particles. The top panel in Figure 5.12 shows the average mass power spectrum from the three strips, each containing 125,000 biased particles (filled circles) and the line shows the average power spectrum from the three strips, with each strip containing 1 million mass

particles. Both of these sets of points have the same radial selection function shape as the 2dF QSO Survey, although they contain many more points. The power spectrum of the biased particles for this test is estimated from many more particles than there will be in the 2dF QSO Survey to reduce the effects of shot noise in the measurement of the power spectrum. Imprinting the $N(z)$ distribution on the biased particles has little effect on the measured power spectrum. This is because all particles at all redshifts were biased to have the same clustering amplitude. However, imprinting the $N(z)$ distribution onto the mass particles increases the amplitude of the mass power spectrum by a factor of $\sim$1.2. This is because there are proportionately more mass particles selected from the $1 < z < 1.5$ range than from the $z > 2$ range and the mass particles at $z \sim 1.25$ have a higher clustering amplitude than those at $z \sim 2.25$.

The lower panel shows the power spectrum of the biased particles divided by the power spectrum of the mass. Over a wide range of scales, there is a linear bias between the two power spectra. The bias should be linear on all scales, apart from the smallest scales where the clustering is non-linear. However, on scales larger than $400h^{-1}$Mpc, the linear bias appears to be a less good approximation. This is just because the two power spectra are not well measured on these scales, due to the effects of the window function and also because the power spectrum is not well sampled on these scales as we embed the survey into a box that is $4{,}000h^{-1}$Mpc. On scales smaller than $60h^{-1}$Mpc the linear bias also appears to be a less good approximation. This is the scale where the FFT no longer gives reliable measurements of the clustering. Fewer particles contribute to the power spectrum of the biased particles than is the case for the mass power spectrum so the power spectrum of the biased particles may be more affected by shot noise and by aliasing due to the FFT more than the mass power spectrum.

5.4.5 Sparse Sampling and the Direct Fourier Transform

So far we have only considered power spectra measured from samples with a higher number density than will be obtained with the 2dF QSO Survey. However, the 2dF QSO survey will contain approximately 25,000 QSOs, observed over two strips. We sparse sample the biased catalogues, as described in Section 5.2.4, so that each strip contains $\sim$12,500 particles, and measure the power spectrum again. Due to the sparsity of QSOs, we must improve the estimation of the power spectrum. There are two ways in which this can be done. First, rather than binning the power spectrum in linear bins, the power spectrum is binned in logarithmic bins of size $\log(k/h\mathrm{Mpc}^{-1})$=0.075. This increases the size of the

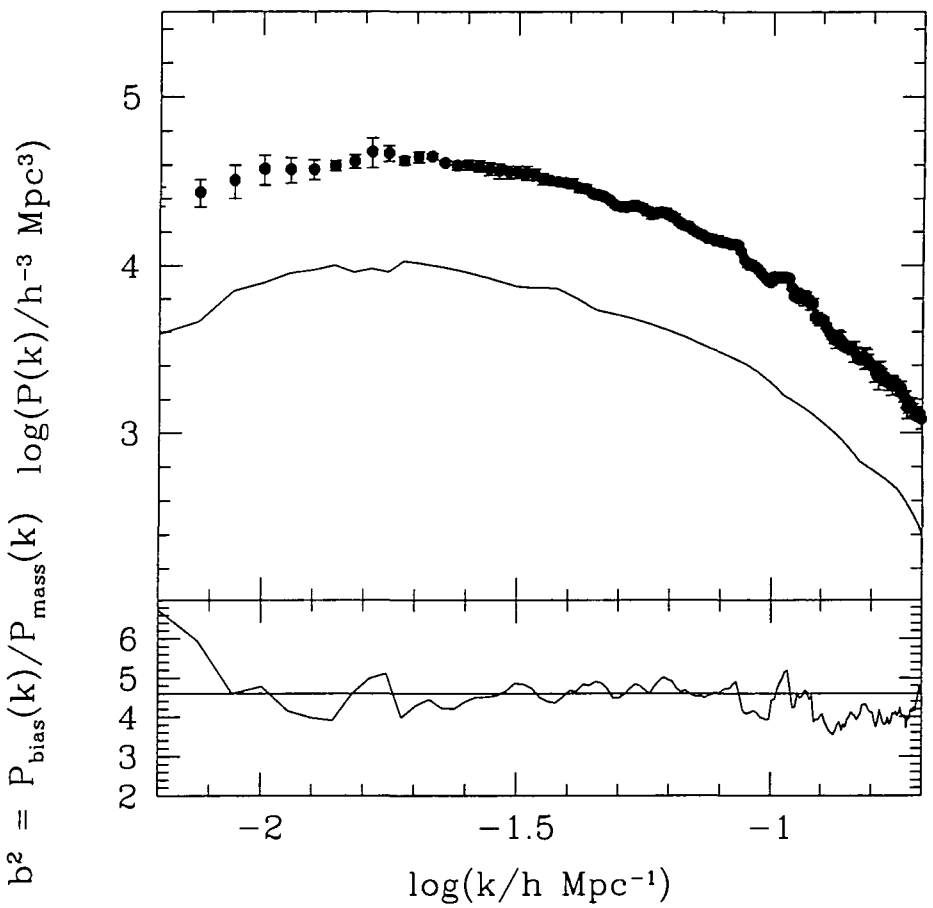

Figure 5.12: The upper panel shows the power spectrum of 125,000 biased particles in redshift space (solid circles) and the line shows the power spectrum of 1 million dark matter particles in redshift space. Both of these sets of points have the same angular and radial selection function as the QSO in the 2dF QSO Survey. The errors come from the dispersion over the three strips, normalised to two strips. The lower panel shows the result of dividing the biased power spectrum by the mass power spectrum. By creating biased catalogues that have constant clustering with redshift the shape of the mass power spectrum has been preserved, with the biased power spectrum simply having a higher clustering amplitude. In this case, the bias parameter, b, has a value of $\sim \sqrt{4.6}$.

bins, particularly on small scales, so more modes contribute to the measurement of the power spectrum. Therefore, the recovered power spectrum is less noisy and the points less correlated. The second improvement is to calculate the power spectrum on small scales, $r \lesssim 100h^{-1}$Mpc, using a direct Fourier transform of the data points. This method is computationally intensive as the time required to calculate the power spectrum increases as N^2, however as there are only 12,500 biased particles on each strip, this method can now be used.

The reason for using the direct Fourier transform on small scales is because the Fast Fourier Transform (FFT) is only accurate on scales larger than the Nyquist frequency, as discussed in Chapter 3 and above in Section 5.4.3. Due to the large volume of the QSO survey, even with the largest FFT grid size possible on the computers available (256), the Nyquist frequency is $(2\pi\ 256)/(2l_{\rm box})$ which corresponds to $\log(k/h{\rm Mpc}^{-1})$=-0.7 ($r \sim 30h^{-1}$Mpc) for a box large enough to encompass the whole 2dF QSO survey. We use the nearest grid point scheme to assign QSOs to the FFT grid, so aliasing can affect the shape of the power spectrum on scales somewhat larger than that corresponding to the Nyquist frequency. We therefore use the direct method to measure the power spectrum on scales smaller than $\sim 100h^{-1}$Mpc to ensure that aliasing will not affect the shape of the power spectrum.

The direct method is calculated by defining

$$\delta_{\mathbf{k}} = \frac{1}{(2\pi)^3} \sum_{j=1}^{N} e^{i\mathbf{k}.\mathbf{x}_j}. \tag{5.10}$$

The definition of the power spectrum is then

$$P(k) = \frac{\int_{k_1}^{k_2} |\delta_k|^2 k^2 \sin\theta {\rm d}\theta {\rm d}\phi {\rm d}k}{\frac{4\pi}{3}(k_2^3 - k_1^3)} \tag{5.11}$$

where k_1 and k_2 are the lower and upper limits of the bin in k-space.

The filled circles in Figure 5.13 show the average power spectrum measured from the three mock catalogues, each containing 12,500 biased particles and with the radial and angular selection function matching that expected in the completed 2dF QSO survey. The errors come from the dispersion over the three mock catalogues and are normalised to two strips to show the size of the errors expected from the 2dF QSO Survey. On scales larger than $100h^{-1}$Mpc ($\log(k/h{\rm Mpc}^{-1})$=-1.2), the FFT was used to measure the power spectrum. On smaller scales, the direct method was used. The scale at which the method changes is marked by the arrow. The open circles show the power spectrum estimated by the FFT down to smaller scales. The effects of aliasing can clearly be seen. The solid

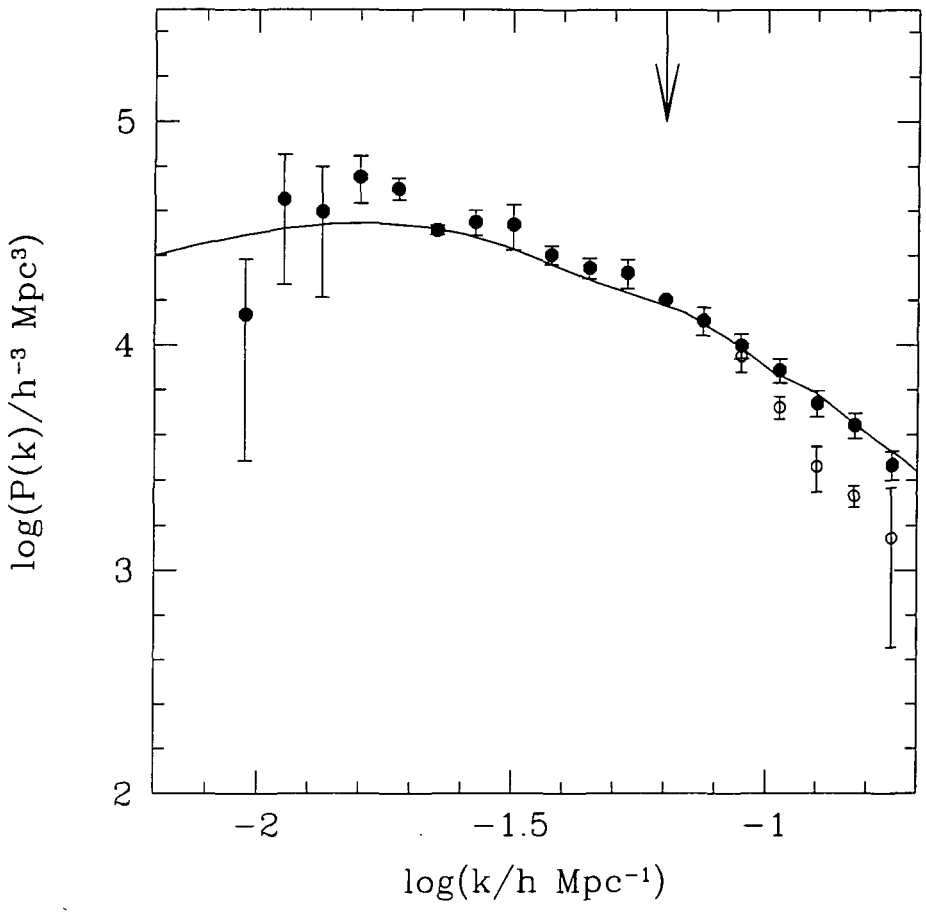

Figure 5.13: The solid line shows the input power spectrum to the simulation, scaled to match the amplitude of the mock catalogue power spectrum. The circles show the average power spectrum of the three mock catalogues in redshift space which have the same angular and radial selection function, as well as the same number of mock QSOs expected in the final 2dF QSO catalogue. The errors are from the dispersion over the three mock catalogues, normalised to two strips. On scales larger than $100h^{-1}$Mpc the FFT was used to calculate the power spectrum whilst the direct FT method was used on scales smaller than this. The scale where the method of estimating the power spectrum changes is shown by the arrow. The open circles show the power spectrum estimated using the FFT. This clearly shows the effects of aliasing, however, by using the direct FT we are able to recover the power spectrum to smaller scales. We are able to recover the shape of the input power spectrum over the range of scales $40 < r < 400h^{-1}$Mpc. On scales larger than $400h^{-1}$Mpc, the errors become large but the power spectrum does turn over at the scale expected in this ΛCDM cosmology.

line in Figure 5.13 shows the input power spectrum to the *Hubble Volume* simulation, scaled to match the power spectrum of the mock catalogues. The agreement between the two power spectra is good. By using two different methods, the power spectrum will be measurable over at least a decade of scale, from $\sim 40h^{-1}$Mpc to $\sim 400h^{-1}$Mpc and possibly out to larger scales, although the power spectrum will have a slightly distorted shape due to the effects of the window function on scales larger than $400h^{-1}$Mpc.

5.4.6 Power Spectrum Errors

So far, only the dispersion over the three mock catalogues has been used to estimate the errors on the power spectrum. These errors agreed with other estimations of the error of the correlation function to within a factor of $\sim$2 but they varied significantly with scale. If many mock catalogues were available, these errors would be more stable and should provide an adequate estimate of the cosmic variance, but with only 3 mock catalogues they have to be used with caution.

Another method for estimating the errors has been suggested by Feldman, Kaiser & Peacock (1994, FKP). Simplifying equation 2.3.2 of FKP for a sample with constant number density as a function of r and, therefore, a weighting of $w = 1$ given to all QSOs gives

$$\frac{\sigma^2(k)}{P^2(k)} = \frac{(2\pi)^3[1 + \frac{1}{\bar{n}P(k)}]^2}{V_k V_s} \tag{5.12}$$

where V_k is the volume of each shell in $k-$space and V_s is the volume of the survey. Every QSO is given an equal weight of $w = 1$ as the survey is essentially a volume limited survey due to the almost flat radial selection function of QSOs in the 2dF QSO Survey (see Figure 5.4(b)).

The volume of the shell in k-space can be estimated in two ways. One is to assume that the shell is spherical and estimate V_k by $4\pi/3(k_2^3 - k_1^3)$ where k_1 and k_2 are the lower and upper limits of the k bin. The other method, which we use here, is to estimate the size of the bin by $V_k = N_k(\Delta k)^3 = N_k(2\pi/l_{\rm box})^3$, where N_k is the number of modes in the k-shell and $(\Delta k)^3$ is the volume of one k-mode. For a sample with high number density, such that the factor $[1+1/(\bar{n}P(k))] \sim 1$, the FKP errors scale as $1/N_k$, similar to Poisson errors.

Figure 5.14 shows a comparison of the errors found using the dispersion over the three mock catalogues (solid line) and the errors of FKP (dashed line). On intermediate scales, $50 \lesssim r \lesssim 200h^{-1}$Mpc the two methods agree within a factor of ~ 2 and indicate that

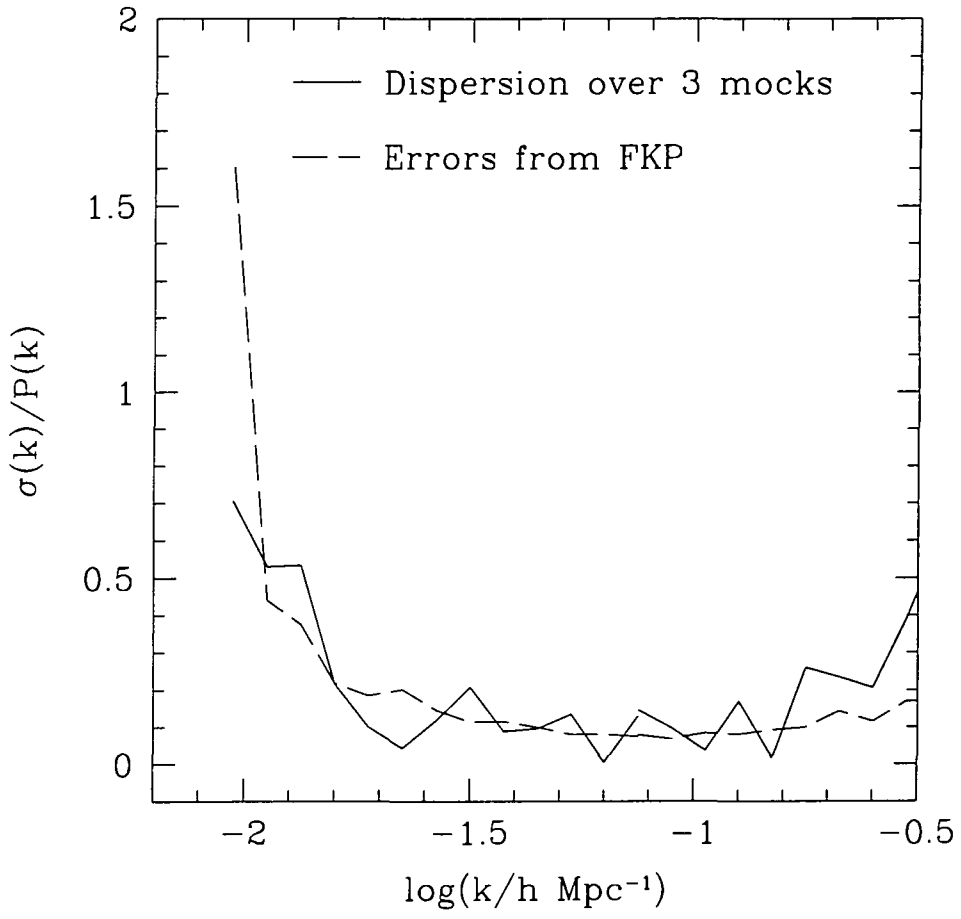

Figure 5.14: Comparing different error estimates for the power spectrum: The solid line shows the fractional error calculated from the dispersion over the three mock catalogues, normalised to two strips. The dashed line shows the errors of Feldman, Kaiser & Peacock (1994), again normalised to two strips. There is reasonable agreement between the two methods over a wide range of scales.

the power spectrum will be measurable to an accuracy of around 10% using the binning adopted here. The FKP errors seem to match the overall trend of the mock catalogue errors over a wide range of scales. On the largest scales, the FKP error is very large as the smallest k-bin is not well sampled when a box of size $l_{\rm box} = 4000h^{-1}$Mpc is used to calculate the FFT, therefore V_k may be underestimated in equation 5.12. At $400h^{-1}$Mpc, where the turnover occurs in the ΛCDM simulation, the errors should still be less than 40%. On the smallest scales, the FKP errors seem to under estimate the errors from the mock catalogues. This is probably because effects, such as edge effects, are not included in the definition of the FKP errors.

5.4.7 Can we measure the turnover?

One of the features we hope to detect with the power spectrum of the 2dF QSO Survey is the turnover expected in CDM power spectra as the position of the turnover contains information on cosmology and can constrain models of structure formation, as discussed in Chapter 2. However, whether we will be able to measure it or not depends on what scale the turnover occurs, if it even occurs at all.

If the cosmology is $\Omega_{\rm m}$ =0.3 and Ω_Λ=0.7, then the power spectrum can be cleanly measured out to $400h^{-1}$Mpc before the Survey geometry starts to affect the recovered shape. If the cosmology is $\Omega_{\rm m}$=1, then the scale where the window function affects the power spectrum is reduced to $300h^{-1}$Mpc.

If no turnover is detected out to $300h^{-1}$Mpc, then this restricts the range of CDM models that could fit the 2dF QSO Survey power spectrum. In Figure 6.10, we show power spectra with a range of different values of the shape parameter, Γ. The standard CDM model has Γ=0.5 and the turnover for this model occurs on scales of $\sim 120h^{-1}$Mpc. This would be detected by the 2dF QSO Survey. In fact, models with $\Gamma \gtrsim 0.25$ have turnovers that would be detected by the power spectrum measured out to $300h^{-1}$Mpc. Therefore, if no turnover is seen in the power spectrum of the 2dF QSO Survey on scales $\lesssim 300h^{-1}$Mpc, the shape parameter must be $\lesssim 0.25$.

However, for a power spectrum to agree with both the constraints from COBE (Smoot et al. 1992) and the present day abundance of clusters (Eke, Cole & Frenk 1996), discussed in Chapter 2, a CDM power spectrum has to have a value of Γ ∼0.2. This is the value adopted for most ΛCDM and τCDM simulations. If Γ=0.2, then the scale of the turnover is $\sim 350h^{-1}$Mpc. If the cosmology is $\Omega_{\rm m}$=1, then the 2dF QSO Survey will struggle to reliably measure a turnover on these scales. However, if the cosmology is $\Omega_{\rm m}$=0.3 and

Ω_Λ=0.7 then the 2dF QSO Survey should just be able to detect it.

5.5 Summary

We have biased dark matter particles from the Hubble Volume simulation to approximately reproduce the clustering found from the first 9,000 QSOs observed in the 2dF QSO survey. Unless these QSOs are found in extremely low or high density regions, we expect that the mock catalogues will be realistic representations of the final survey. We assume that the clustering amplitude is constant with redshift, consistent with the work of Croom & Shanks (1996) and early results from the 2dF QSO survey (Croom et al. 2000). We are able to reproduce this by changing the biasing parameters with redshift.

We have tested three different correlation function estimators (Peebles 1980, Hamilton 1993, Landy & Szalay 1993) and find that the correlation function does not depend on which estimator is used. Even the simple estimator of Peebles (1980), equation 5.5, agrees with the estimators of Hamilton (1993) and Landy & Szalay (1993) to within 10%. In order to reduce the chance of introducing systematic effects into the estimation of the correlation function, we adopt the method of Hamilton (1993). We have also tested different error estimators. Errors found from the dispersion over the three mock catalogues vary from scale to scale, due to the small number of mock catalogues, and must be treated with caution on any one scale. We find that Poisson errors are a reasonable estimate of the errors on scales out to $40h^{-1}$Mpc but on scales larger than this, they underestimate the error. On scales larger than $40h^{-1}$Mpc we choose constant N_q errors. We predict that the correlation function will be measureable out to $1,000h^{-1}$Mpc scales and over the range $2.5 \lesssim r \lesssim 40h^{-1}$Mpc the errors should be less than 20%, using the binning adopted here and assuming that the mock catalogues have similar clustering to the final 2dF QSO Survey and $\Omega_{\rm m}$=0.3, Ω_Λ=0.7 for the cosmology. On scales larger than this, the correlation function tends to zero so we quote the true error rather than the fractional error. If $\xi = 0$, then constant N_q errors give an absolute value for the error to be 0.0063, a factor of 5 improvement on previous surveys.

Accurate measurements of the volume averaged correlation function in at least four redshift bins should be possible from the 2dF QSO Survey. If the clustering measured from the 2dF QSOs does have a constant clustering amplitude, we predict the errors on the volume averaged correlation function with $r < 20h^{-1}$Mpc to be around 10%. For assumed values of the cosmological parameters, $\Omega_{\rm m}$ and Ω_Λ, the relative QSO-mass bias

as a function of redshift can be calculated. This may then be used to discriminate between models of QSO biasing, such as those of Matarrese et al. (1997) and Colín et al. (1999).

Measurements of the power spectrum of QSO clustering will be possible from the 2dF QSO Survey. Assuming Ω_m=0.3 and Ω_Λ=0.7, the power spectrum should be free from the effects of the window function out to scales of $400h^{-1}$Mpc. The distortion introduced on scales $400 \lesssim r \lesssim 500h^{-1}$Mpc is fairly small. By using a combination of the FFT and the direct Fourier transform, we are able to measure the power spectrum reliably over a decade in scale, from 40 to 400 h^{-1}Mpc. The direct Fourier Transform is required on small scales as aliasing occurs in the power spectrum measured by the FFT on scales smaller than $\sim 60h^{-1}$. This is due to the large volume of the 2dF QSO survey and the maximum FFT grid size possible with present computing facilities. Over the range of scales $50 \lesssim r \lesssim 200h^{-1}$Mpc, the errors from the dispersion over the mock catalogues and the errors from the method of Feldman, Kaiser & Peacock (1994) seem to agree fairly well and we predict that the errors will be ~10% using the binning adopted here, again assuming that the mock catalogues have similar clustering to the final 2dF QSO Survey.

Whether or not we will measure the turnover in the power spectrum depends on what scale the turnover occurs at. The power spectrum should be cleanly measured out to $300h^{-1}$Mpc if the cosmology is Ω_m=1 and $400h^{-1}$Mpc if the cosmology is Ω_m=0.3, Ω_Λ=0.7. This means that if no turnover is measured in the power spectrum, Γ has to be less than 0.25. If Γ=0.2, the value which agrees with both COBE and cluster normalisation, then the turnover occurs on scales of $\sim 350h^{-1}$Mpc. If Ω_m=1 is the cosmology, the 2dF QSO Survey will struggle to detect it but if Ω_m=0.3, Ω_Λ=0.7 it should be detected.

Chapter 6

Power Spectrum Results from the January 2000 2dF QSO Catalogue

6.1 Introduction

The power spectrum is now established as one of the favoured methods of quantifying clustering (for examples of power spectrum analysis from the surveys considered in this Thesis see Chapter 3 as well as Baugh & Efstathiou 1993, Tadros & Efstathiou 1996, Lin et al. 1996). As discussed in previous Chapters, there are many advantages in measuring the power spectrum, the power spectrum is only affected by uncertainties in the mean density on the largest scales (Cole, Fisher & Weinberg 1995) and, if objects trace the distribution of the mass in the Universe in a simple manner, meaningful comparisons between the mass power spectrum from models of structure formation and the object power spectrum can be made.

The power spectrum of optically selected QSOs has never been measured. As discussed in Chapter 2, QSOs are detected out to high redshifts and so QSO surveys have a large volume. However, existing surveys only cover small areas of sky and contain only a small number of QSOs. The survey geometry and low QSO number density within the surveys would have severely restricted the range of scales over which the power spectrum could have been reliably measured. The completed 2dF QSO Survey will provide a factor of $\sim$2 improvement over existing surveys in terms of the area of sky surveyed and a factor of $\sim$25 improvement in the number of QSOs in the survey.

In Chapter 5, we predicted that the 2dF QSO power spectrum would be measurable and that the effects of the window function would be negligible on scales $40 \lesssim r \lesssim 400h^{-1}$Mpc. This prediction was for the completed survey, assuming $\Omega_{\rm m}$=0.3 and Ω_Λ=0.7 for the cosmology. The current incompleteness of the survey will restrict the range of

scales over which the power spectrum can be measured, due both to the smaller number of QSOs and, more importantly, to the complicated geometry of the incomplete survey. The smallest dimension (the declination direction) will limit the range of scales over which the power spectrum can be measured in the completed survey and this problem will be exacerbated further here. However, the QSOs that have been observed have not been selected from the input catalogue by their colour or magnitude and so are observed over the full depth of the 2dF QSO Survey. The January 2000 catalogue therefore covers a large volume.

In this Chapter, we apply power spectrum analysis to the ~9,000 QSOs that have so far been observed in the 2dF QSO survey. Measurements of the luminosity function (Boyle et al. 2000) and the two point correlation function (Croom et al. 2000) have been made. Already the vast improvement of the 2dF QSO survey over existing QSO surveys in terms of the errors, the scales over which the correlation function can be measured and the number of redshift bins that both the correlation function and the luminosity function can be measured in, has been demonstrated. We describe the current data set in Section 6.2 and look at the angular distribution of the QSOs in order to create the random catalogues needed in the power spectrum analysis. In Section 6.3 we briefly describe the method of power spectrum analysis. In Section 6.4 we present our results by comparing the QSO power spectrum to the mock catalogues, created in Chapter 5. We compare the QSO power spectrum to power spectra of present day galaxies and clusters and we see how the QSO power spectrum evolves with redshift. In Section 6.5 we compare the QSO power spectrum with power spectra of models of large scale structure and in Section 6.6 we draw our conclusions.

6.2 The January 2000 Data Set

As of January 2000, the 2dF QSO survey is approximately 40% completed. 18,779 objects have been observed using the 2dF instrument on the AAT. From these objects, ~ 9,000 have been positively identified as QSOs by an automated procedure (Miller et al. 2000). For this analysis, we only consider QSOs with redshifts in the range $0.3 < z < 2.4$ where the completeness is expected to be better than ~90% (Boyle et al. 2000). This gives a sample of 8608 QSOs.

Observations have been made over two 5×75 degree strips in the North and South Galactic Pole regions. The South Galactic Pole strip is centred at $\delta = -30°$, with 21^h40^m

$\lesssim \alpha \lesssim 03^h15^m$ while the North Galactic Pole strip is centred at $\delta = 0°$ with $9^h50^m \lesssim \alpha \lesssim 14^h50^m$. We will refer to these regions as the NGP and SGP. There are 3709 QSOs in the NGP region and 4899 QSOs in the SGP region. The $N(z)$ distribution of all the observed QSOs with redshifts in the range $0.3 < z < 2.4$ is shown in Figure 5.4. A low order polynomial was fitted to this distribution so that the radial selection function of the QSOs could be imprinted on the mock catalogues (see Section 5.2.4). In Figure 6.1(a), we show the $N(z)$ distribution of the QSOs in the NGP (dashed histogram) and SGP (solid histogram) with redshifts in the same range. The line corresponding to each histogram shows the polynomial fit scaled to match the number of QSOs in each strip. This Figure shows that the form of the $N(z)$ distribution found from all the QSOs, used in Chapter 5 in the construction of the mock QSO and random catalogues, is consistent with the $N(z)$ distribution of QSOs on each strip. This means the radial selection function of the mock and random catalogues constructed in Chapter 5 matches that of the observations on each strip.

For each QSO, we have an angular position and a redshift. In order to convert from a redshift into a distance, we need to adopt a cosmology. To compare the power spectrum with the predictions from Chapter 5 we assume a cosmology with $\Omega_m = 0.3$ and Ω_Λ=0.7 which matches the cosmology of the *Hubble Volume* simulation. We also consider an $\Omega_m = 1$ cosmology. Other cosmologies, such as open cosmologies, could also be considered but here we restrict our analysis to flat cosmologies, consistent with the recent results from balloon experiments such as Boomerang (de Bernardis et al. 2000) and Maxima (Balbi et al. 2000).

Figure 6.1(b) shows the radial selection function of QSOs in the NGP (dashed line) and SGP (solid line) assuming $\Omega_m = 0.3$ and Ω_Λ=0.7. The radial selection function only varies slowly as a function of scale for QSOs on both strips. The value of $\bar{n}(r)$ changes by only a factor of $\sim$3 over the whole range of scales, whereas the value of $\bar{n}(r)$ for galaxies in the Durham/UKST Survey changes by a factor of $\sim$ 1000 over a far smaller range of scales (Figure 2.3). If Ω_m=1 is assumed, there is also little variation in the QSO $\bar{n}(r)$ as a function of scale.

6.2.1 Constructing the Random Catalogue

Observations have been made across both the NGP and SGP strips. An optimal tiling algorithm for the 2dF Galaxy Redshift survey was developed by the 2dF Galaxy Redshift team to allow as many galaxies and QSOs as possible to be observed in each field. This

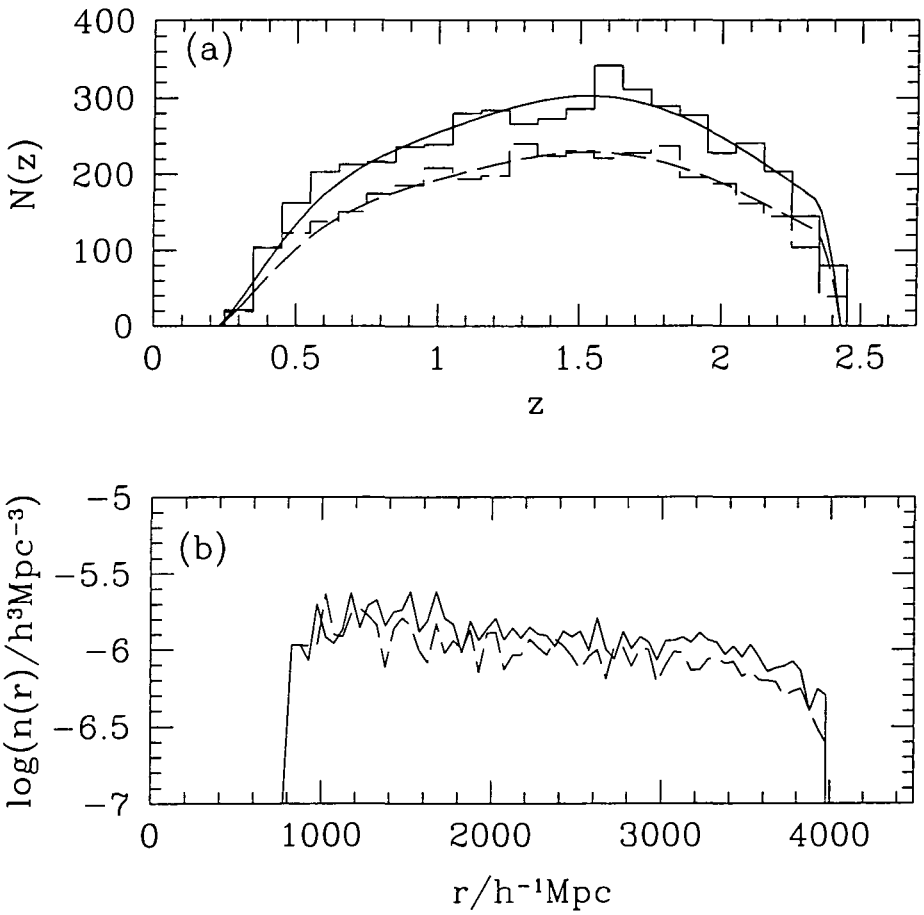

Figure 6.1: Panel (a) shows the number of QSOs per 0.1 bin of redshift in the SGP (solid histogram) and the NGP (dashed histogram). The lines show the polynomial fit to the $N(z)$ distribution of all the QSOs observed so far with redshifts in the range $0.3< z <2.4$, used in Chapter 5. The polynomial has been normalised to match the number of QSOs in each strip. Panel (b) shows the radial number density of QSOs in bins of $\Delta r = 50h^{-1}$Mpc in the SGP (solid line) and the NGP (dashed line) calculated assuming $\Omega_m = 0.3$ and Ω_Λ=0.7. The number density varies slowly as a function of scale, decreasing only by a factor of ~ 3 over the whole range.

means that the survey currently has a patchy angular selection function which has to be matched by the random catalogue.

In order to generate the random catalogue, a completeness map for the NGP and SGP has to be constructed. The objects that will be observed and those that have been observed are divided up into bins of 1 degree in RA by 0.5 degrees in dec. The fraction of QSOs already observed in each field is then simply defined as

$$F_{\rm rel} = \frac{\text{Number of objects observed on each field}}{\text{Total number of objects contained in each field}}. \quad (6.1)$$

We use a simple definition for the fraction of QSOs observed on each field which can easily be updated as more and more QSOs are observed. If one field has a higher stellar contamination than another, this should be compensated for in the above definition for $F_{\rm rel}$. Ideally, we would like to know the ratio of observed QSOs to the number of QSOs in each field but, in advance of the observations, we do not know which objects in the input catalogue are QSOs. The completeness maps for the NGP and SGP are shown in Figure 6.2. Note that observations in the NGP have been made fairly randomly over the whole sky, whereas in the SGP, they have been preferentially made in the more northerly regions. However, there are more bins that have lower completeness in the NGP than in the SGP. In the NGP there are currently 156 bins that have $0.01 < F_{\rm rel} < 0.1$, whereas in the SGP, there are only 72 bins with $0.01 < F_{\rm rel} < 0.1$ (out of ~ 560 bins on each strip that have been partly observed).

Next, we need to construct a random catalogue for the NGP and SGP regions. To construct this, we follow the method described in Chapter 3 and Chapter 5. We generate random points within a box that is larger than the strip under consideration. Points that have redshifts outside of the range $0.3 < z < 2.4$ are immediately rejected, as are particles with coordinates that do not over lap with the geometry of the strip under consideration. The selection function of the QSOs is imprinted onto the random catalogue, as described in Chapter 5, before the incompleteness map is imprinted onto the random catalogue. For each point in the random catalogue, we calculate which 1×0.5 deg^2 bin the point lies in. We generate a random number between 0 and 1 and if this value is less than $F_{\rm rel}$ for that bin, the particle remains in the random catalogue. Otherwise the particle is rejected as it lies in part of the survey that has not been fully observed. The random catalogue contains many points and has the same angular and radial selection function as the completed strip is expected to have but otherwise contains no clustering.

The power spectrum of the window function of the NGP and SGP regions is shown in

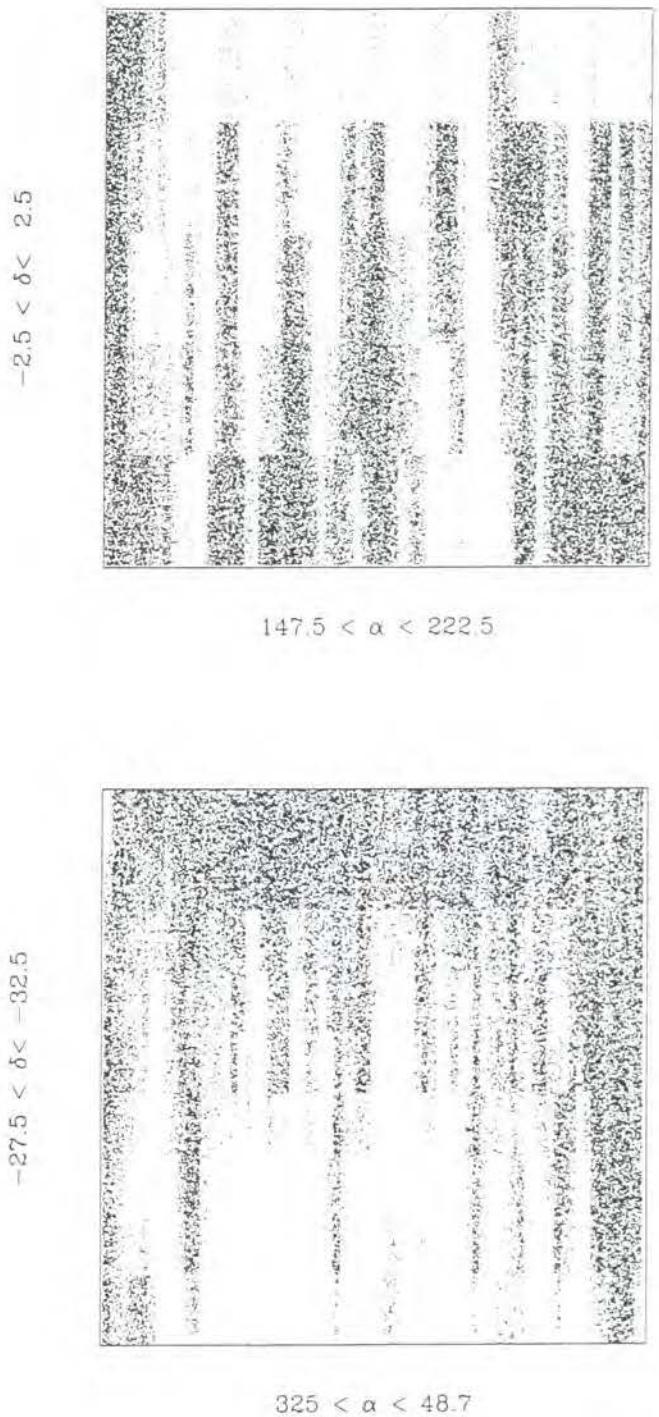

Figure 6.2: The completeness map for the NGP (top panel) and the SGP (bottom panel). The scales of the dec axis has been exaggerated relative to the RA range. The regions that are heavily shaded are more complete than the areas with lighter shading and there are some areas that have not yet been targeted.

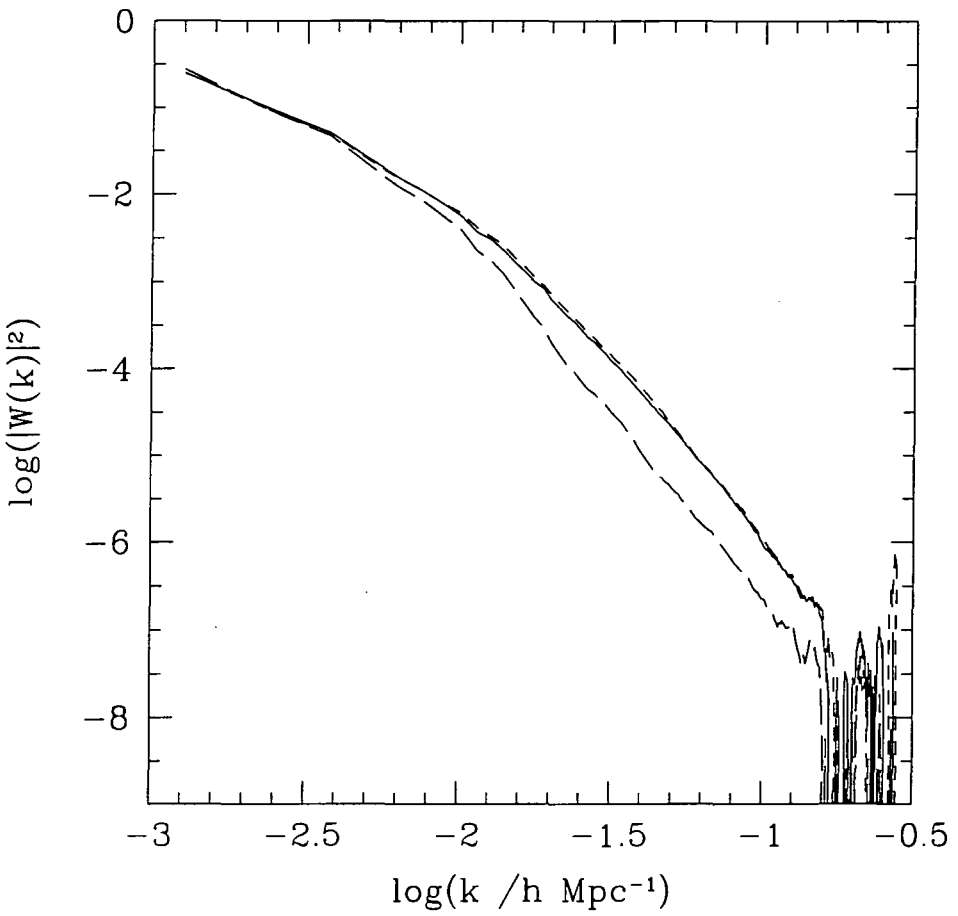

Figure 6.3: The window function of the current 2dF QSO survey in the NGP (short dashed line) and the SGP (solid line) compared to the window function of a completed $5° \times 75°$ strip. These are calculated assuming $\Omega_m = 0.3$, Ω_Λ=0.7. The window function of the NGP and SGP is a steep power law out to scales of $\sim 400h^{-1}$Mpc as opposed to $600h^{-1}$Mpc when the survey is completed.

Figure 6.3. The short dashed line shows the window function of the NGP and the solid line shows the window function of the SGP. The long dashed line shows a prediction for the window function of the survey when it is fully completed, i.e. when QSOs are measured over the full 5×75 degree strip, as calculated in Chapter 5. The current window function of the NGP is very similar to that of the SGP. The SGP currently contains slightly more QSOs and covers a slightly larger area than the NGP but the QSOs in the SGP are observed preferentially in the more northerly part of the strip. The power spectrum of the window function from the NGP and SGP is not quite as steep as the power spectrum of the window function from the completed strip, nor does the power law part extend to the same large scales that the window function of the completed survey does. The power spectrum from the NGP and SGP is only unaffected by the window function on scales $r \lesssim 300h^{-1}$Mpc as opposed to $\sim 400h^{-1}$Mpc for the completed survey, assuming Ω_m=0.3, Ω_Λ=0.7, found by convolving the power spectrum of the window function with the input power spectrum, as discussed in Chapter 5. The current window function affects the power spectrum on scales of $\gtrsim 200h^{-1}$Mpc if Ω_m=1 is assumed.

6.3 Measuring P(k)

The method used to measure the power spectrum is the same as that described in Chapter 5 and is similar to the method used in Chapter 3. We use a FFT to obtain the power spectrum on the largest scales and a direct method of computing the Fourier transform on scales smaller than $r < 100h^{-1}$Mpc ($\log(k/h\mathrm{Mpc}^{-1})$=-1.2). We bin the Fourier modes up in logarithmic bins of size $\delta\log(k/h\mathrm{Mpc}^{-1}) = 0.075$. We use a box size of 4,000 h^{-1}Mpc and a grid size of 256 when computing the FFT.

The radial selection function of the QSOs is shown in Figure 6.1(b); $\bar{n}(r)$ only varies slowly as a function of scale for QSOs in the SGP (solid line) and for QSOs in the NGP (dashed line). This means that all QSOs can be given equal weight when estimating the power spectrum (see Chapters 3 and 5 as well as Feldman, Kaiser & Peacock (1994)).

Two different cosmologies are considered to convert the QSO redshifts into distances as discussed in Section 6.2. These are Ω_m=0.3, Ω_Λ=0.7 (referred to as the Λ cosmology hereafter) and Ω_m=1 (EdS hereafter). The power spectrum of the QSOs in the NGP and SGP are averaged together, weighted by the inverse of the error on each scale, in order to obtain a single QSO power spectrum for the 2dF QSO survey for each cosmology.

6.3.1 Error Determination

The errors on the power spectrum are estimated in two ways. The first is to use the ΛCDM *Hubble Volume* lightcone simulation and imprint the completeness map of the NGP and the SGP on to the mock catalogues, constructed in Chapter 5. The errors are then the dispersion found over the three realisations of the power spectra. As there are only three mock catalogues, this is not ideal as on any single scale the error may differ from other error estimators by up to a factor of ~2 (see Figure 5.14). The second way in which the errors are estimated is to use the method of Feldman, Kaiser & Peacock (1994), assuming that all QSOs carry equal weight when estimating the QSO power spectrum. This method is described in more detail in Chapter 5.

The values of the power spectrum assuming either cosmology (Λ or EdS) alongside the errors estimated from the mock catalogues and the method of Feldman, Kaiser and Peacock (1994, FKP) are given in Table 6.1. The mock catalogue errors on the power spectrum assuming the EdS cosmology are found by assuming that the fractional error on the power spectrum estimated from the EdS cosmology at k is the same as the fractional error estimated from the Λ cosmology at k' with $k = k'/(f_{\perp}^2 f_{\parallel})^{1/3}$. The terms $f_{\perp}$ and $f_{\parallel}$ are discussed in more detail in Chapter 7 and are calculated at the median redshift of the survey, $\bar{z}$=1.4. To first order, they allow a power spectrum calculated in one cosmology to be translated into another cosmology. We have also obtained mock catalogue errors by assuming that the ΛCDM *Hubble Volume* lightcone simulation has a cosmology of Ω_m=1. The mock catalogue errors are then found from the dispersion over the three power spectra. The two methods yield similar errors. A lightcone simulation with Ω_m=1 is really required to obtain more reliable errors and, for both cosmologies, more than three mock catalogues are required. We will therefore show FKP errors for the power spectra in this Chapter, although on the smallest scales these errors may underestimate the true error as effects such as edge effects are not fully included in the FKP errors.

6.4 Results

6.4.1 Comparison with the Mock Catalogues

The points in Figure 6.4 show the power spectrum of QSOs in the NGP (a) and the power spectrum of QSOs in the SGP (b). The Λ cosmology is assumed to allow a comparison with the mock catalogues constructed in Chapter 5. The solid line in both panels shows

the average power spectrum from the mock catalogues with the incompleteness maps imprinted on them. Errors are from the method of FKP. Over the range $-1.7 \lesssim \log(k/h$ Mpc$^{-1}) \lesssim -0.9$ there is reasonable agreement between the prediction from the simulation and the measured QSO power spectrum.

On the very largest and smallest scales, the power spectrum from the SGP strip seems to be slightly higher than the power spectrum from the mock catalogues. The excess power is currently not that significant, the QSO power spectrum on small scales is $\sim 2\sigma$ higher than the mock catalogue power spectrum and on large scales the excess power seen in the SGP power spectrum is consistent with the power spectrum from the mock catalogues, due to the large errors. If the excess power in the SGP power spectrum is still detected when the Survey is completed it will warrant further attention.

We combine the power spectra from the NGP and SGP, weighted by the inverse of the FKP variance, to obtain a single power spectrum, assuming the Λ cosmology. We combine the power spectra from the NGP and SGP strip found from the mock catalogues in a similar manner. These two power spectra are shown in Figure 6.4(c). The points show the QSO power spectra with FKP errors and the solid line shows the simulation power spectra with the current incompleteness included. The dashed lines show the 1σ FKP errors on the redshift space power spectrum from the mock catalogues. As seen in the NGP and SGP strips, there is a good match with the mock catalogue power spectra over a wide range of scales with perhaps evidence that the QSO power spectrum has a slightly steeper slope than the mock catalogues from the ΛCDM *Hubble Volume* simulation predicts.

In Figure 6.5, we compare the QSO power spectrum to the input power spectra of the two *Hubble Volume* simulations, discussed in Chapter 2. Panel (a) compares the QSO power spectra measured assuming the Λ cosmology to the real space ΛCDM input power spectrum. This power spectrum has a shape of $\Gamma \sim 0.17$. In Panel (b), we compare the QSO power spectrum, assuming the EdS cosmology, with the real space τCDM input power spectrum. This has a shape of Γ=0.21. In both cases, we have normalised the input power spectra to match the QSO power spectra as closely as possible to allow a comparison between the two shapes to be more readily made. The QSO power spectrum in panel (a) seems to have a slightly higher amplitude than the input power spectrum at $\sim 150h^{-1}$Mpc (log $(k/h\text{Mpc}^{-1})$ = -1.4). On smaller and larger scales, the QSO power spectrum matches the input power spectrum reasonably, though the slope could still be slightly steeper than predicted by ΛCDM. The QSO power spectrum in panel (b) seems to

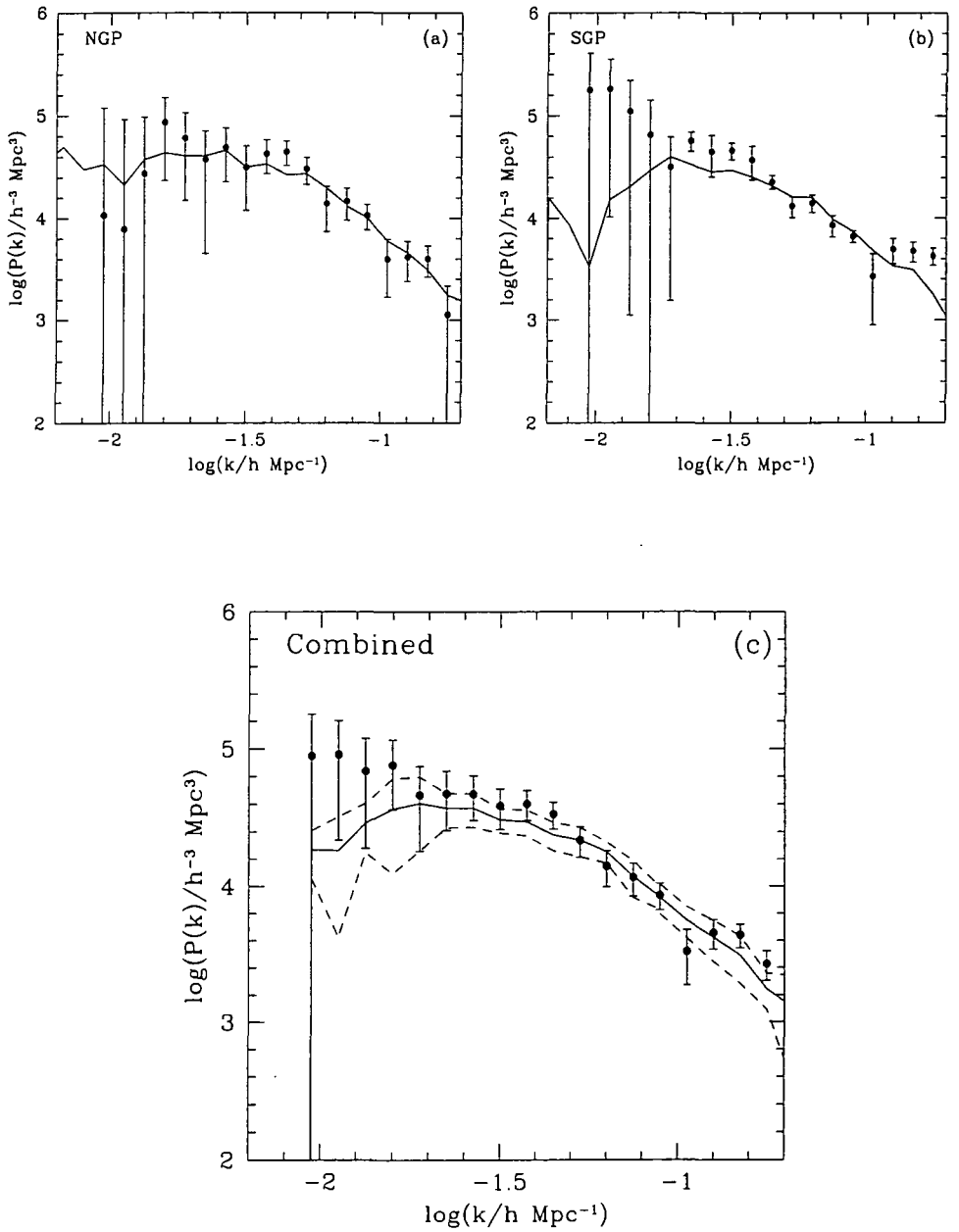

Figure 6.4: The points in (a) show the power spectrum of the QSOs in the NGP, the points in (b) show the power spectrum of QSOs in the SGP and the points in (c) show the QSO power spectrum, found by combining the NGP and SGP power spectra, see text. The solid line in all cases is the redshift space prediction from mock catalogues (constructed in Chapter 5) with the relevant incompleteness map imprinted on them. These two power spectra are combined to give the prediction in (c). The errorbars show the FKP errors on the QSOs, combined for panel (c). The dashed lines in (c) show the FKP errors predicted from the mock catalogues. We use a FFT to measure the power spectra on scales larger than $100h^{-1}$Mpc ($\log(k/h\mathrm{Mpc}^{-1})$=-1.2) and the direct method on smaller scales. The Λ cosmology is assumed.

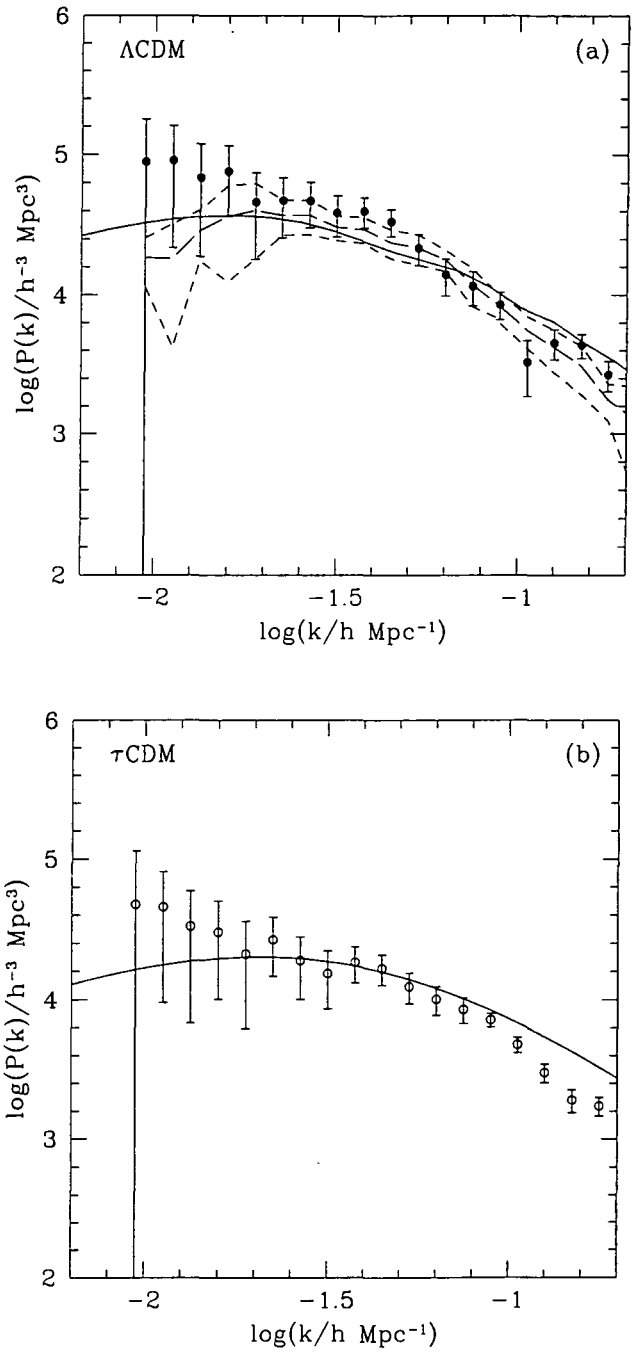

Figure 6.5: The points in each panel show the power spectra from the January 2000 catalogue with the Λ cosmology assumed (a) and the EdS cosmology assumed (b). The solid lines show the input power spectra to the *Hubble Volume* simulations. The input power spectrum to the ΛCDM simulation is shown in (a) and the input power spectrum to the τCDM simulation is shown in (b). The input power spectra are normalised to match the QSO power spectra as closely as possible to allow the shapes to be readily compared. The dashed lines in (a) show the mock catalogue redshift space power spectrum and errors from Figure 6.4(c), discussed in the text.

Cosmology	$\log(k/h\mathrm{Mpc}^{-1})$	$\mathrm{P}(k/h^{-3}\mathrm{Mpc}^3)$	FKP Error	Mock Error
$\Omega_\mathrm{m} = 0.3\ \Omega_\Lambda = 0.7$	-2.025	88850.0	89264.4	107378.
	-1.950	91165.0	79364.8	97146.0
	-1.875	68878.5	49923.4	54809.7
	-1.800	76131.5	39691.3	51050.1
	-1.725	46253.5	28213.8	23625.3
	-1.650	47372.0	21651.8	11925.8
	-1.575	47078.0	16758.6	13186.7
	-1.500	38778.0	12576.9	11221.7
	-1.425	40053.5	9777.84	8744.66
	-1.350	33716.0	7429.43	10643.1
	-1.275	21793.0	5520.01	6477.37
	-1.200	14037.5	4138.92	2477.88
	-1.125	11583.6	3173.32	3042.91
	-1.050	8543.00	1901.79	643.436
	-0.975	3310.92	1438.52	446.310
	-0.900	4537.82	1114.19	856.628
	-0.825	4366.90	859.449	782.356
	-0.750	2680.26	659.315	445.418
$\Omega_\mathrm{m} = 1$	-2.025	47406.0	66670.8	57392.7
	-1.950	45476.6	35961.7	40360.3
	-1.875	33168.9	26337.3	30753.6
	-1.800	30182.7	20077.5	31770.9
	-1.725	21093.9	14888.5	17476.4
	-1.650	26636.5	11979.1	8506.65
	-1.575	18929.0	8891.52	5251.78
	-1.500	15341.5	6721.81	3214.05
	-1.425	18428.5	5279.51	3979.39
	-1.350	16660.5	4037.18	4518.70
	-1.275	12376.0	3036.37	2701.72
	-1.200	10062.3	2319.46	1776.59
	-1.125	8539.40	1779.54	2361.60
	-1.050	7209.50	811.235	542.673
	-0.975	4788.00	611.806	646.008
	-0.900	3006.79	463.963	534.719
	-0.825	1907.08	354.378	359.812
	-0.750	1732.72	272.931	195.514

Table 6.1: The 2dF QSO power spectrum for two different assumptions about the underlying cosmology and with two different estimates of the errors. The power spectrum for each cosmology is found by averaging together the power spectrum of the NGP and SGP strips, weighting by the inverse of the errors. The errors are found by combining the errors from the NGP and SGP strips.

also have a somewhat steeper slope than the input power spectrum to the τCDM *Hubble Volume* simulation.

The dashed line shows the mock catalogue power spectrum, as shown in Figure 6.4. Over the range of scales $60 < r < 300h^{-1}$Mpc, the mock catalogue power spectrum matches the input power spectrum to within 1σ, although the mock catalogue power spectrum has a slightly steeper slope than the input power spectrum. Figure 5.13, however, shows the mock catalogue power spectrum from the completed survey. This matches the input power spectrum over a wider range of scales and the slopes seem to be more similar. This suggests that the current incompleteness is having some systematic affect on the shape of the mock catalogue power spectra which may also affect the power spectrum measured from the January 2000 catalogue, although until the 2dF QSO Survey is complete, we cannot test this directly. Instead, in Section 6.5 we limit the range of scales over which we compare the current QSO power spectrum to models of structure formation, such that the power spectrum, measured from the mock catalogues with the incompleteness included, matches the shape of the input power spectrum of the simulation to within 1σ.

The errors on the completed survey are predicted to be a factor of $\sim$1.5 smaller on intermediate scales ($100 \lesssim r \lesssim 250h^{-1}$Mpc), where the power spectrum from the NGP and SGP can already be measured fairly accurately. This is because the error on the power spectrum scales approximately proportional to $1/\sqrt{V_s}$ and approximately 2/5 of the Survey has been observed (see equation 5.12 in Chapter 5). On larger and smaller scales, the errors on the power spectrum of QSOs should be more than a factor of 1.5 times larger than those predicted from the completed survey. This is due to the effects of the geometry on the largest scales and effects of the smaller number of QSOs on small scales.

6.4.2 Comparison with P(k) from Galaxy and Cluster Surveys

As already seen in Chapter 3, the power spectrum of galaxies has been measured from several different galaxy redshift surveys. Measurements of the galaxy power spectrum from optically selected surveys are so far only able to determine the power spectrum on scales out to $\sim 100h^{-1}$Mpc. In Figure 6.6, we compare the power spectrum estimated from the QSOs currently observed in the 2dF QSO Survey, assuming two different cosmologies, with the power spectrum from the Durham/UKST survey (see Chapter 3). The lines show the QSO power spectrum, estimated assuming the Λ cosmology (solid line) and

the EdS cosmology (dashed line). The solid points show the power spectrum from the Durham/UKST survey (circles - flux limited with P=$8000h^{-3}$Mpc^3, triangles - volume limited with $z_{\rm max} = 0.06$).

Figure 6.6 shows that the amplitudes of the QSO and galaxy power spectra are consistent. The QSO power spectrum, computed assuming the Λ cosmology (solid line), has a slightly lower amplitude than the galaxy power spectra but apart from the dip at $60h^{-1}$Mpc ($\log(k/h\mathrm{Mpc}^{-1})$=-1) the two power spectra agree to within 1-2σ. If an EdS cosmology is assumed, the amplitude of the QSO power spectrum (dashed line) is approximately a factor of 2 lower than the galaxy power spectra. Croom et al. (2000) also finds better agreement between the amplitude of the galaxy and QSO correlation function if the Λ cosmology is assumed.

We also compare the QSO power spectra to the real space APM galaxy power spectrum of Baugh & Efstathiou (1993), inferred from measurements of the angular correlation function $w(\theta)$. The lines in Figure 6.7 are the QSO power spectra, assuming the two different cosmologies as discussed above. The squares show the APM power spectrum. On large scales, the difference between the galaxy redshift space power spectrum and the galaxy real space power spectrum should just be a shift in amplitude due to bulk motions of galaxies causing an increase in the redshift space power spectrum amplitude (Kaiser 1987). In Chapter 3, we measured a value of β=0.60$\pm$0.35 by comparing the shapes of the Durham/UKST and APM power spectra. Translating the APM real space power spectrum into redshift space would boost the amplitude by a factor of ~ 1.4. On scales out to 100 h^{-1}Mpc, there is reasonable agreement between the amplitudes of the different power spectra. On scales larger than $100h^{-1}$Mpc, the APM power spectrum flattens off, whereas the QSO power spectra continue to rise. More large scale power was found in the APM power spectrum than in the standard Cold Dark Matter model, discussed in Chapter 2. The QSO power spectrum has even more large scale power than the APM power spectrum and may have even more power than the ΛCDM models.

A final comparison is made between the QSO power spectra and the power spectrum of rich clusters of galaxies. In Figure 6.8, the lines are as in Figure 6.6 and the points show the APM Survey cluster power spectra taken from Tadros, Efstathiou & Dalton (1998). The cluster power spectra are found assuming $\Omega_{\rm m}$=1 (filled circles) and $\Omega_{\rm m}$=0.2, Ω_Λ=0.8 (open circles). As the clusters are only observed out to a redshift of $\sim$0.2, the effects of cosmology on the clustering amplitude are small so the two cluster power spectra have a similar amplitude and shape. As can clearly be seen in Figure 6.8, the QSOs have a

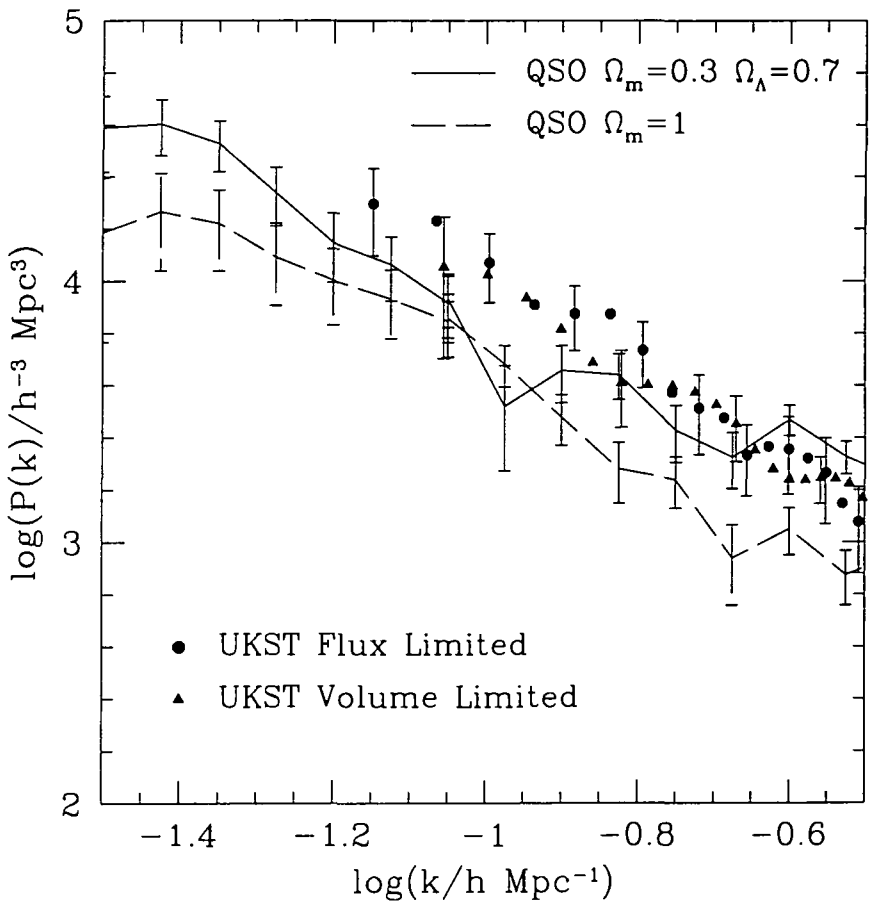

Figure 6.6: A comparison of galaxy and QSO power spectra. The lines show the QSO power spectra with the Λ cosmology (solid line) and the EdS cosmology (dashed line) assumed. The solid symbols are from the Durham/UKST survey, Chapter 3. The circles are power spectra estimated from a flux limited sample and the triangles are power spectra estimated from a volume limited sample. The errors on the QSO power spectra are FKP errors and the errors on the galaxy power spectra, shown only on every other flux limited point for clarity, are from mock catalogues.

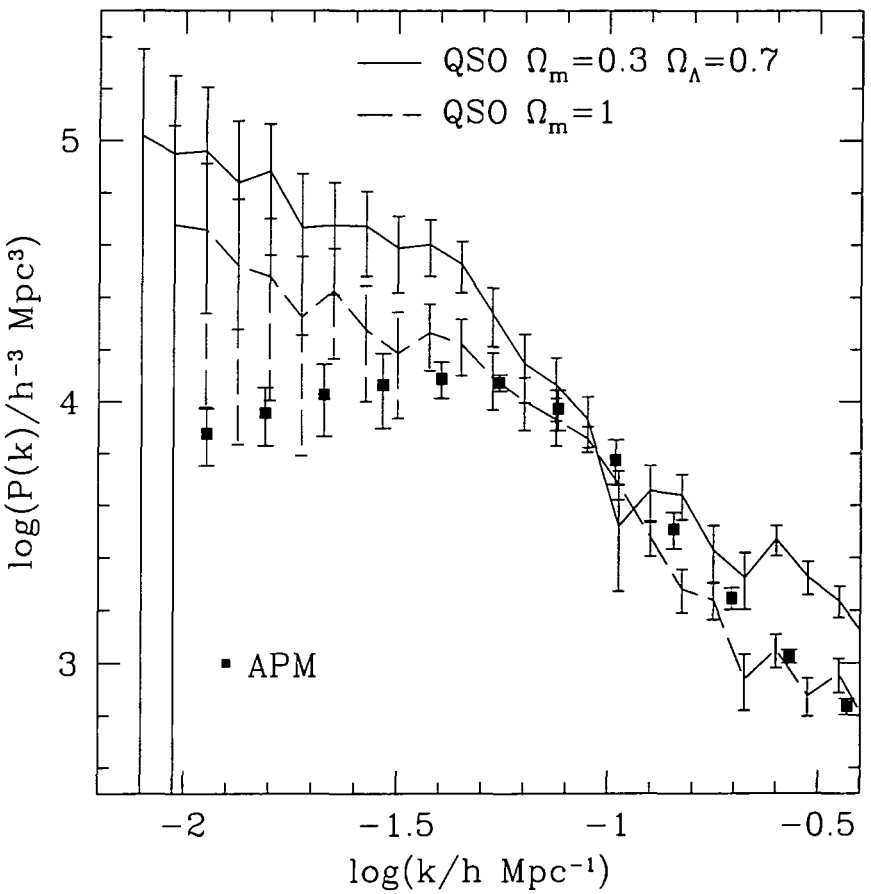

Figure 6.7: A comparison between redshift and real space power spectra. The lines show the QSO power spectra with the Λ cosmology (solid line) and the EdS cosmology (dashed line) assumed. The squares show the APM real space galaxy power spectrum taken from Baugh & Efstathiou (1993). The errors on the QSO power spectra are FKP errors. The errors on the APM power spectrum are obtained by measuring the scatter in power spectra obtained from four separate regions of the APM survey.

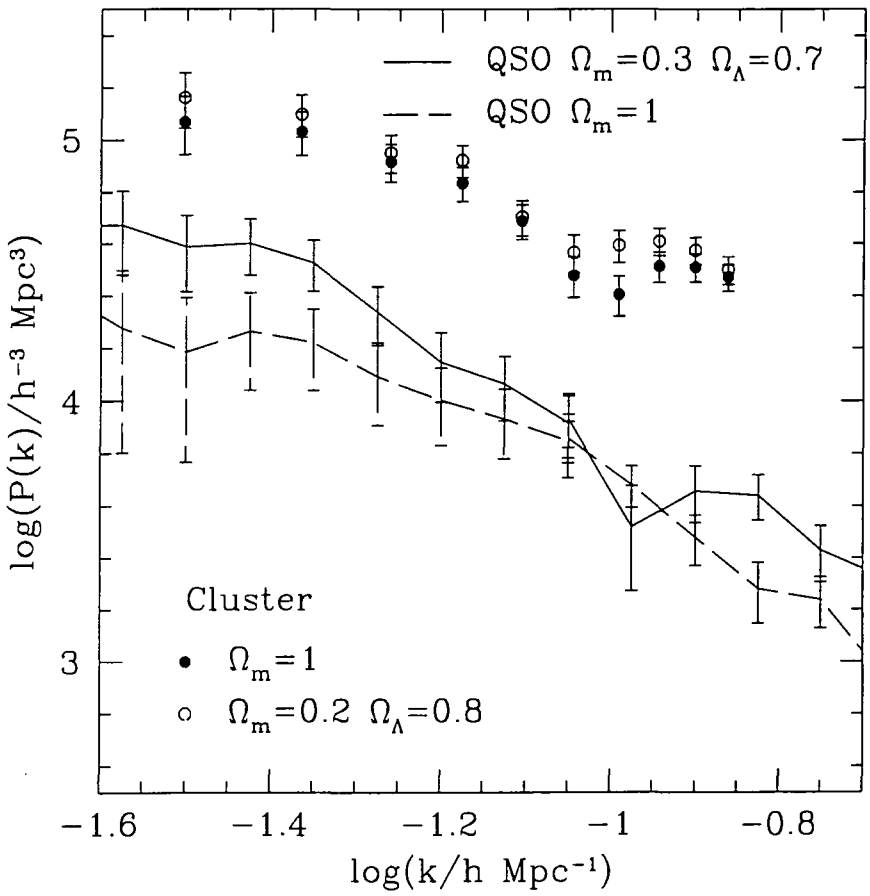

Figure 6.8: A comparison between cluster and QSO power spectra. The lines show the QSO power spectra with the Λ cosmology (solid line) and the EdS cosmology (dashed line) assumed. The points show the power spectrum of rich clusters with $\Omega_m = 1$ (filled circles) and Ω_m=0.2, Ω_Λ=0.8 (open circles) taken from Tadros, Efstathiou & Dalton (1998). The errors on the QSO power spectra are FKP errors and the errors on the cluster power spectra come from mock catalogues.

lower clustering amplitude than the present day rich clusters. The relative bias between the two types of power spectra is $b_{\rm rel} \sim \sqrt{5}$ for the Λ cosmology and $b_{\rm rel} \sim \sqrt{6}$ for the EdS cosmology.

The galaxy and cluster power spectra are essentially independent of the assumed cosmology. However, the QSO power spectrum does depend on the cosmology assumed to measure the power spectrum due to the higher median redshift of this sample. If we assume that the galaxy-mass bias, the cluster-mass bias and the QSO-mass bias are independent of scale so that the bias does not change the shapes of the different power spectra, we can apply a simple cosmology test to the QSOs by comparing the slopes of the QSO power spectra, calculated with different assumptions of the cosmology, and the galaxy and cluster power spectra.

The cluster power spectrum has an approximate slope of -1.13±0.15 over the range of scales $50 < r < 200h^{-1}$Mpc whereas the QSO power spectrum with the Λ cosmology assumed has a slope of ∼-1.70±0.20 and the slope assuming the EdS cosmology is ∼-1.22±0.15. The flux limited galaxy power spectrum from the Durham/UKST Survey has a slope of ∼-1.90±0.15 over the range of scales $20 < r < 100h^{-1}$Mpc. Over the same range of scales, the QSO power spectrum, assuming the EdS cosmology, has a similar slope of ∼-1.80±0.20 whereas the QSO power spectrum has a slope of ∼-1.40±0.20, assuming the Λ cosmology. We find that the QSO power spectrum calculated assuming the EdS cosmology matches the slope of the galaxy and cluster power spectra better than the QSO power spectrum assuming the Λ cosmology. This is tentative evidence that the EdS cosmology is closer to the true underlying cosmology than the Λ cosmology, if the assumption that bias is independent of scale holds for galaxies, clusters and QSOs.

Power spectra from the current generation of galaxy surveys and the QSO power spectrum from the January 2000 catalogue cannot be measured accurately enough for strong results on cosmology to be made from the slopes of the power spectra. The slopes of the power spectra are sensitive to features, such as the dip at $60h^{-1}$Mpc in the Λ cosmology power spectrum and the QSO power spectrum from the January 2000 catalogue may appear more steeply sloped than the QSO power spectrum from the completed survey, due to incompleteness effects. A stronger test may be possible from the completed 2dF QSO and Galaxy Redshift Surveys when the errors on the power spectrum will be reduced and the shapes of two power spectra can be compared over a wider range of scales.

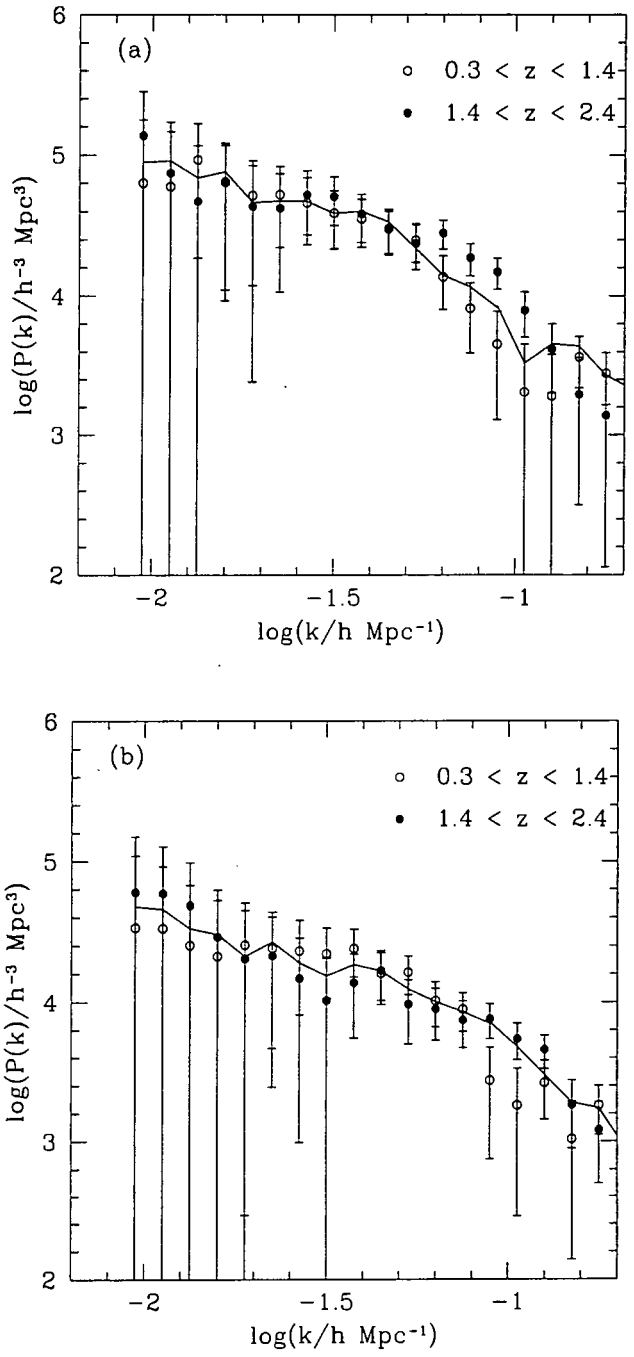

Figure 6.9: The power spectrum of QSOs measured at different redshifts. The Λ cosmology is assumed in (a) and the EdS cosmology is assumed in (b). In both panels, the solid line shows the power spectrum estimated from all the currently observed QSOs, as in Figure 6.5. The open circles show the power spectrum of QSOs with redshifts in the range $0.3 < z < 1.4$ and the filled circles show the power spectrum of QSOs with redshifts in the range $1.4 < z < 2.4$. The errors are FKP errors.

6.4.3 The Evolution of QSO Clustering

To determine how QSO clustering evolves as a function of redshift, we split the QSOs in each strip of the survey into two redshift bins containing roughly equal numbers of QSOs, one with QSOs in the range $0.3 < z < 1.4$ ($\bar{z} = 1.0$) and the other with $1.4 < z < 2.4$ ($\bar{z} = 1.8$). The power spectrum is then measured from each subsample on each strip, assuming the two cosmologies discussed in Section 6.2. The results from the NGP and SGP for each redshift bin and for each cosmology are averaged together as before. All the power spectra in Figure 6.9(a) are estimated assuming the Λ cosmology, whilst in panel (b), EdS is assumed. In both panels, the solid line shows the power spectrum of all the QSOs, the open circles show the power spectrum of the low redshift QSOs and the filled circles show the power spectrum of the high redshift QSOs.

When the QSOs are split up into the two samples, the errors on the measured power spectra are increased. However, on scales around 100-200h^{-1}Mpc the power spectrum can still be fairly well measured, the fractional errors are around 40% if either FKP errors or errors from the mock catalogue are used. There seems to be reasonable agreement between the power spectra measured from the two redshift bins over a wide range of scales. If the Λ cosmology is assumed, then at $\log(k/h\mathrm{Mpc}^{-1}) \sim -1.2$ the high redshift power spectrum has a $\sim 2\sigma$ higher amplitude than the low redshift power spectrum. However, the total power from the two redshift bins over the range $-1.7 < \log(k/h\mathrm{Mpc}^{-1}) < -1$ agrees to within 1σ. This suggests that QSO clustering does not evolve strongly with redshift.

This result is consistent with the work of Croom et al. (2000). Using the volume averaged correlation function, Croom et al. (2000) have measured the clustering from the January 2000 catalogue in five bins of redshift. If the EdS cosmology is assumed then little evolution is seen in the QSO clustering, consistent with Croom & Shanks (1996). If the Λ cosmology is assumed, slow evolution is seen in the QSO clustering, with QSOs at high redshift having a slightly higher clustering amplitude than QSOs at low redshift. The value of $\bar{\xi}(r < 20h^{-1}\mathrm{Mpc})$ found in each of the redshift bins for each assumed cosmology agrees with the value found for galaxies ($\bar{\xi}(r < 20h^{-1}\mathrm{Mpc})$=0.30 assuming a power law with $r_0 = 6h^{-1}\mathrm{Mpc}$ and γ=-1.7) to within $\sim 2\sigma$.

At lower redshifts, there is further evidence that suggests that QSO clustering does not evolve strongly as a function of redshift as QSO clustering at low redshift is also consistent with that of galaxies. Georgantopoulos & Shanks (1994) find that low-redshift, IRAS selected Seyferts have a low clustering amplitude of $\bar{\xi}(r < 10h^{-1}\mathrm{Mpc})$=0.5±0.31

(which assuming γ=-1.7 translates to a value of $\bar{\xi}(r < 20h^{-1}\mathrm{Mpc})$ =0.15). These objects were IRAS selected and may have a slightly lower clustering amplitude than optically selected QSOs, as seen for galaxies (Chapter 3). However, Smith, Boyle & Maddox (1995, 1999) also find that radio-quiet QSOs with redshifts in the range $0 < z < 0.7$ are found in similar environments as galaxies. 90% of optically selected QSOs are expected to be radio-quiet QSOs. As the galaxies and QSOs are found in similar clustering environments, they should have similar clustering properties. Smith, Boyle & Maddox (1995) find that the QSO-galaxy angular cross correlation function is indistinguishable from the APM galaxy-galaxy angular correlation function.

The slow evolution in QSO clustering was interpreted previously in terms of cosmology (see, for example, Croom & Shanks 1996). The slow evolution either suggests that Ω_m is fairly high, $\gtrsim 0.3$, and that bias evolves as a function of redshift to counteract the decrease in amplitude of the dark matter clustering. The alternative is that Ω_m is low, less than 0.1, as then the dark matter clustering does not evolve strongly with redshift. Determining cosmology from the evolution of QSO clustering alone is not possible due to the degeneracy between cosmology and bias, although in Chapter 7 we show that the QSO-mass bias can help constrain cosmology when the results are combined with other techniques for constraining cosmology from the 2dF QSO Survey.

If Ω_m is high, then a model for the evolution of the QSO-mass bias is required. Croom & Shanks (1996) describe such a model in which QSOs form in density peaks above a certain threshold. This threshold is fairly low such that QSOs are not highly biased objects. However, one possible disadvantage of this model is that no use is made of the relationship between QSO space density and the absolute amplitude of QSO clustering. In fact, to match the observed number density of QSOs, QSOs may have to sparse sample the density peaks as an assumption is made in this model that QSOs are long lived. One basis for this assumption is the simple interpretation of the pure luminosity evolution observed for the QSO luminosity function over the range $0.35 < z < 2.3$ (Boyle et al. 2000).

More recently, other authors have looked at bias models which make direct predictions for the clustering amplitude of QSOs, either at the average redshift of the survey or as a function of redshift. The models of Haiman & Hui (2000) and Martini & Weinberg (2000) predict the clustering amplitude of QSOs as a function of QSO lifetime. If QSOs have a long life time then they are predicted to live in rare objects and therefore the QSO clustering amplitude would be high (higher than the clustering amplitude of present day galaxies). This is not seen, the QSO power spectrum at $\bar{z} = 1.4$ has a similar amplitude

to that of galaxies at present day. This would suggest that QSOs are short lived, although the assumption that long lived QSOs are found in the rarest, most highly biased objects may not be correct. Long lived QSOs may sparse sample density peaks above a low threshold.

The models of Matarrese et al. (1997) and Moscardini et al. (1998) predict the evolution of QSO clustering by setting a threshold for the halo mass within which QSOs reside. At lower redshifts, massive halos become more abundant and less biased so the QSO clustering decreases with redshift in these models. Croom et al. (2000) finds models with a minimum halo mass of $10^{12}M_{\odot}$ to be consistent to the clustering seen in the volume averaged correlation function as a function of redshift from the January 2000 catalogue. However, the threshold above which QSOs may form may not be constant with redshift.

Until more is known about QSO physics and, in particular, QSO formation, it is hard to discriminate between the different models. The only conclusion that may perhaps be made is that QSOs do not seem to be extremely biased objects as the clustering of the QSO sample as a whole is consistent with that of present day galaxies and the slow evolution of QSOs and results from the environments of QSOs suggest that at low redshift, QSOs may also have a similar clustering amplitude as galaxies.

6.5 Comparison with Models of Large Scale Structure

As discussed previously, it is the power spectrum of the mass density field that is predicted by models of large scale structure. If the bias between the dark matter and QSOs can be written as $P(k)_{QSO} = b^2 P(k)_{mass}$, with b a constant (which should be the case on large scales (Coles 1993, Cole et al. 1998, Mann, Peacock & Heavens 1998)), it is possible to compare meaningfully the shapes of mass and QSO power spectra.

The power spectrum of QSOs is measured over a lightcone. However, Figure 5.11 in Chapter 5 shows that the shape of the power spectrum of the mass estimated from the lightcone is very similar to the shape of the power spectrum of the mass measured at the median redshift. Therefore we compare the shape of the QSO power spectrum measured here with model power spectra calculated at the median redshift, $z = 1.4$, of the 2dF QSO Survey.

The power spectrum from the 2dF QSO Survey is measured in redshift space, whereas the model power spectra are calculated in real space. Small scale peculiar velocities affect the shape of the redshift power spectrum on small scales. By comparing the shapes

of the redshift space correlation function and the correlation function inferred from the Fourier Transform of the input power spectrum, measure in real space (Figure 5.1), we find that the small scale peculiar velocities have a negligible effect on the shape of the power spectrum on scales $\gtrsim 5h^{-1}$Mpc.

On large scales, the effect of redshift space distortions is to boost the amplitude of the power spectrum according to

$$P_s(k) = P_r(k)\left(1 + \frac{2}{3}\beta + \frac{1}{5}\beta^2\right), \tag{6.2}$$

(Kaiser 1987) with $\beta = \Omega_{\rm m}^{0.6}/b$ and r and s indicating the real space and redshift space power spectra. The redshift space distortions are caused by bulk motions of QSOs. If the bias is scale independent the redshift space power spectrum just has a higher amplitude than the real space power spectrum but the shapes should be consistent.

However, we find that the current incompleteness may be introducing a bias into the shape of the QSO power spectrum. The shape of the power spectrum from the mock catalogues with the incompleteness imprinted on them is slightly steeper than the input power spectrum to the simulation. On scales down to $60h^{-1}$Mpc, the mock catalogue power spectrum with the incompleteness included matches the input power spectrum to within 1σ but on smaller scales, the mock catalogue power spectrum steepens and the disagreement becomes more significant, see Figure 6.5(a). Therefore, we fit the model power spectra to the QSO power spectrum, assuming the Λ cosmology, down to $60h^{-1}$Mpc only. We have no mock catalogues with the EdS cosmology to test if incompleteness affects the shape of the EdS power spectrum from the January 2000 catalogue. However, $60h^{-1}$Mpc in a Λ cosmology roughly translates to $40h^{-1}$Mpc in an EdS cosmology at $z = 1.4$ so we will limit the fit to that scale.

We set the limit on large scales to be the scale where the geometry of the survey affects the shape of the power spectrum. This corresponds to $200h^{-1}$Mpc in the EdS cosmology and $300h^{-1}$Mpc in the Λ cosmology.

We compare the QSO power spectra to model CDM power spectra with different values of Γ, where $\Gamma = \Omega_{\rm m}h$, as described in Chapter 2. We assume the fit to the CDM transfer function given by Bardeen et al. (1986). The power spectrum is calculated at $z = 1.4$ rather than at $z = 0$ to approximate the dark matter power spectrum averaged over the lightcone. We assume a value of σ_8 consistent with the cluster normalisation for each cosmology, as given by Eke, Cole & Frenk (1996). For each cosmology, the model power spectra are χ^2 fitted to the QSO power spectrum to find the factor, $b_{\rm eff}$, required

Ω_{CDM}	Ω_Λ	Γ	b_{eff}	σ rejection	Line Type (Figure 6.10)
0.3	0.7	0.5	2.96	$>4\sigma$	solid
0.3	0.7	0.4	2.67	3.8σ	short - long
0.3	0.7	0.3	2.51	3.0σ	short dashed
0.3	0.7	0.2	2.22	2.1σ	long dashed
0.3	0.7	0.1	1.93	$< 1\sigma$	dotted
1	0	0.5	5.43	$> 4\sigma$	solid
1	0	0.4	5.05	3.5σ	short - long dashed
1	0	0.3	4.62	2.9σ	short dashed
1	0	0.2	4.10	2.2σ	long dashed
1	0	0.1	3.45	$< 1\sigma$	dotted

Table 6.2: The model power spectra parameters as shown in Figure 6.10 and described in the text. The value of b_{eff} is the value of the bias that best matches the amplitude of the model power spectrum to the QSO power spectrum in each cosmology. The penultimate column gives the level of rejection of the model to the data.

to match the two as well as possible.

We choose values of Γ in the range $0.1 < \Gamma < 0.5$ which more than covers the range of values that fit current galaxy power spectra, such as the APM real space power spectrum (Eisenstein & Zaldarriaga 1999). The models are summarized in Table 6.2. We note that ΛCDM models with $\Gamma \gtrsim 0.3$ are not physical motivated but we are just comparing a range of models with different shapes to the QSO power spectrum and we do indeed find that models with values of $\Gamma \gtrsim 0.3$ give a poor fit to the QSO power spectrum with the Λ cosmology assumed.

In Figure 6.10, we show the QSO power spectrum with the Λ cosmology (solid circles, a) and with the EdS cosmology (open circles, b). The lines have different values of Γ: Γ=0.5 (solid), Γ=0.4 (short - long dashed), Γ=0.3 (short dashed), Γ=0.2 (long dashed) and Γ=0.1 (dotted). Models with Γ=0.5 are ruled out at more than 4σ assuming either cosmology. This is in agreement with results from galaxy surveys. More large scale power was found in the galaxy correlation function from the APM Survey than expected from the standard CDM Γ=0.5 model (Efstathiou, Sutherland & Maddox 1990). This led to variants of the CDM model, such as ΛCDM and τCDM to be developed.

In both cases, the best agreement between the models and the QSO power spectra is found with a model with Γ=0.1. Models with Γ=0.2 are slightly too shallow to match

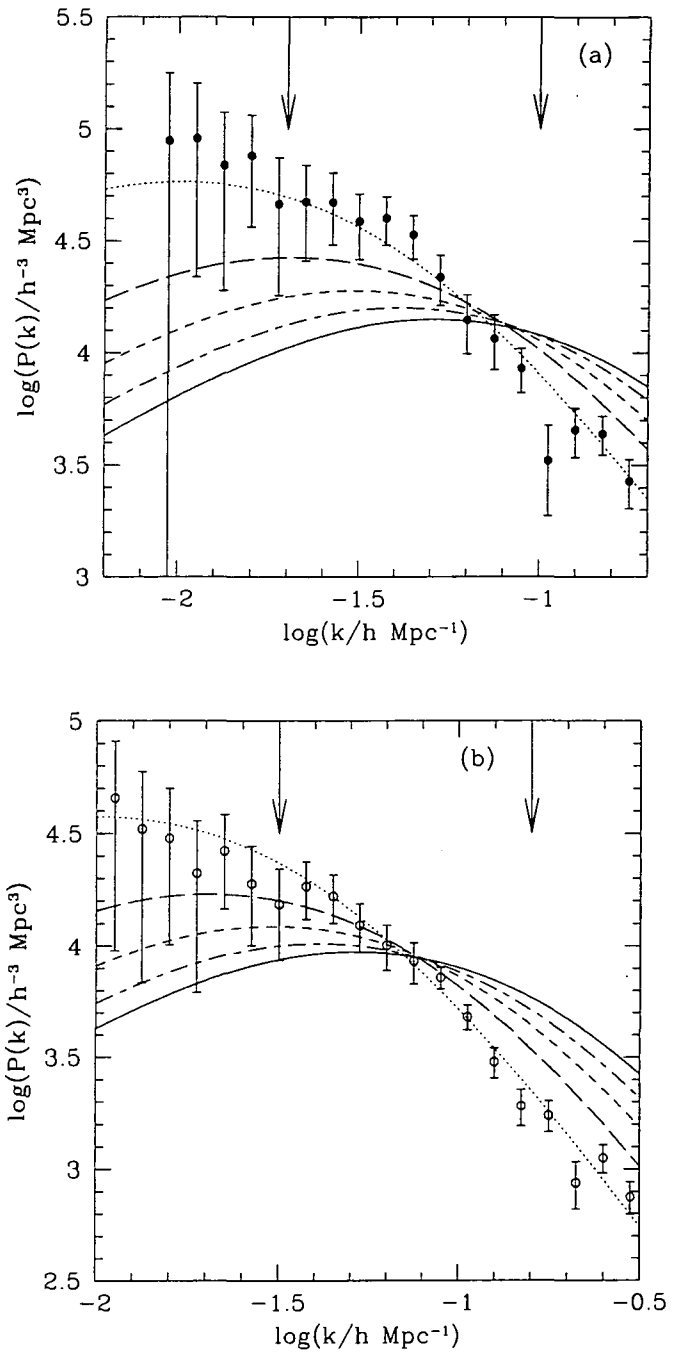

Figure 6.10: Comparison of different model power spectra to the QSO power spectrum in two cosmologies. The solid circles in panel (a) show the QSO P(k) with the Λ cosmology assumed and the open circles in panel (b) show the QSO P(k) with the EdS cosmology assumed. In both panels the matching lines show the same Γ values with Γ=0.5 (solid line), 0.4 (short - long dashed), 0.3 (short dashed), 0.2 (long dashed) and 0.1 (dotted). Each line is plotted at the amplitude that gives the best fit to the QSO power spectrum. The arrows indicate the range of scales over which the fit was made.

the shape of the QSO power spectra and are rejected at the 2σ level. The input power spectrum to the ΛCDM *Hubble Volume* has a shape of Γ=0.17. This is marginally rejected at the 1-2σ level. The τCDM input power spectrum with γ=0.21 is rejected at the 2σ level.

If we extend the range of scales over which we fit the model power spectra to the QSO power spectra to a decade in scales, i.e. down to $30h^{-1}$Mpc in the Λ cosmology and $20h^{-1}$Mpc in the EdS cosmology, then we get even stronger constraints on the models that fit, for example models with Γ=0.2 are rejected in both cases at the 3σ level. However, the results are sensitive to features, such as the dip seen at $60h^{-1}$Mpc in the Λ cosmology, and the effects of incompleteness may be worse on these scales. Once the 2dF QSO Survey is completed, the range of scales over which the fit can be made will be extended at large and small scales and the fitting should be less sensitive to features, such as the dip at $60h^{-1}$Mpc.

The results found here are, however, similar to a those found in Croom et al. (2000). Croom et al. find that models with Γ=0.5 do not provide a good fit to the correlation function measured from the January 2000 catalogue assuming the EdS and the Λ cosmology. Models with Γ=0.2 best fit the correlation function if the EdS cosmology is assumed but models with Γ=0.1 are required to fit the correlation function with the Λ cosmology assumed.

This test is quite simplistic as possible effects of the incompleteness have not been included in the shape of the model power spectrum, although by fitting to a limited range of scales we should have limited the effect that this can have on the models that fit the power spectrum (compare Figure 6.5 and 5.13). In order to thoroughly test models of structure formation, full N-body simulations, such as the *Hubble Volume* are required. However, due to the large volume of the 2dF QSO survey, large amounts of super computing time are required to carry out such simulations so it is currently not possible to run them for a wide range of cosmological models. Physically motivated models for selecting QSOs within the simulation are also required as a linear bias between the mass and the QSOs may be an over simplification, even on these large scales

6.6 Conclusions

We have demonstrated that useful preliminary results on the power spectrum can be obtained from the current status of the 2dF QSO survey. The power spectrum, found

by combining the power spectra of the NGP and SGP strips, is measured to an accuracy of $\lesssim 30\%$ over the range of scales $50 \lesssim r \lesssim 200h^{-1}$Mpc, assuming the Λ cosmology. On these scales, the errors should be approximately 1.5 times larger than the prediction for the completed survey.

The shape and amplitude of the power spectrum is in reasonable agreement with the predicted power spectrum from the *Hubble Volume* simulation, with tentative evidence for more excess power and a steeper spectral slope than that of Λ and τCDM.

The current 2dF QSO survey power spectrum, measured assuming either cosmology, has an amplitude close to that of optically selected galaxies. The range of scales where the two overlap is small, but once the 2dF Galaxy Redshift Survey is completed, an interesting comparison between power spectra of optically selected QSOs and galaxies over a wide range of scales will be possible. The power spectrum of rich APM clusters measured by Tadros, Efstathiou & Dalton (1998) has a far higher clustering amplitude than the QSO power spectrum for both choices of cosmology. The slope of the QSO power spectrum matches the slope of the galaxy and cluster power spectrum slightly better if the EdS cosmology is assumed than if the Λ cosmology is assumed.

We measure the QSO power spectrum in two bins of redshift and find that the clustering can only evolve slowly as a function of redshift as QSOs with a median redshift of $\bar{z}$=1.0 have a clustering amplitude consistent with that of QSOs with a median redshift of $\bar{z}$=1.8. This is consistent with the results of Croom et al. (2000), who finds little evolution in the amplitude of the volume averaged correlation function as a function of redshift if the EdS cosmology is assumed but slow evolution is found if the Λ cosmology is assumed.

We compare the shape of the QSO power spectrum to models of large scale structure over the range of scales $40 < r < 200h^{-1}$Mpc in the EdS cosmology and $60 < r < 300h^{-1}$Mpc in the Λ cosmology. We find that models with $\Gamma \sim 0.1$ are required in order to match the large scale power seen in the QSO power spectrum. However, this result may be biased slightly low due to the effects of incompleteness. Once the survey is completed, we will be able to tell if incompleteness is affecting the shape of the QSO power spectrum from the January 2000 catalogue or if models with Γ=0.1 are required to match the QSO power spectrum.

Chapter 7

Using the 2dF QSO Survey to constrain Λ

7.1 Introduction

One of the challenges that still faces cosmologists is to determine the values of the underlying cosmological parameters of the Universe, such as $\Omega_{\rm m}$, Ω_Λ and Hubble's constant. In this Chapter, we describe a method through which $\Omega_{\rm m}$ and Ω_Λ may be constrained. Hubble's constant is discussed further in Chapter 8.

During the 1990's, it became clear that there was a problem with the favoured Einstein-de Sitter, $\Omega_{\rm m}$=1 model. Hubble's constant was being measured as $H_0 \sim 70$ km s^{-1}Mpc^{-1} which gives an age for the Universe of $\sim$ 10 Gyr if $\Omega_{\rm m}$=1. Globular clusters were estimated to be older than the Universe itself at 16$\pm$2 Gyr (Renzini et al. 1996), although recent estimates of the ages of Globular clusters are slightly lower, e.g. 11.8$\pm$2.1 Gyr (Gratton et al. 1997). However, inflationary theory, in its simplest form, implies a flat universe. To comply with this requirement, a cosmology where $\Omega_{\rm m}+\Omega_\Lambda$=1 was suggested as an alternative to the Einstein-de Sitter model (Peebles 1984). Invoking a cosmological constant, Λ term ($\Omega_\Lambda = \Lambda c^2/(3H_0^2)$), provides a solution to the problem of the age of the Universe, as there is less mass in the Universe to slow down the expansion, resulting in a greater age. A cosmology with non-zero Λ also matches the amplitude of fluctuations on both COBE (Smoot et al. 1992) and cluster scales (Eke, Cole & Frenk 1996), as well as allowing the Universe to remain flat. However, there is no real physical motivation for a Λ term. The Λ term was originally introduced by Einstein into his field equations to create a static Universe. He later referred to this as his 'biggest blunder' after Hubble discovered the expansion of the Universe (Hubble 1929).

One of the most elegant ways of testing for the existence of a non-zero Λ term was put forward by Alcock & Paczyński (1979). They suggested that if the clustering of galaxies or QSOs, parallel and perpendicular to the line of sight, was computed assuming an

Einstein-de Sitter cosmology but the Universe had a different cosmology, for example one which had $\Omega_\Lambda \neq 0$, then a squashing or elongation of the clustering pattern would occur more in one direction than in the other. This is because assuming the wrong cosmology affects the clustering pattern parallel to the line of sight differently to that perpendicular to the line of sight. The cosmology *assumed* for measurements of the clustering can have any value of Ω_m and Ω_Λ but if these values differ from the *underlying cosmology* of the Universe, distortions are introduced into the clustering pattern. The only assumption required is that the clustering in *real space* is on average spherically symmetric.

Clustering statistics measured from galaxy or QSO redshift surveys are, however, not measured in real space and so they have redshift space distortions imprinted on them. On small scales, virialized clusters appear elongated along the line of sight in redshift space. These are the so called 'fingers of God'. They arise because the redshift due to cosmological expansion and redshift due to peculiar motions cannot be separated. The redshift is converted into a distance assuming that there are no peculiar motions, which distorts the shape of the cluster. On large scales, coherent infall squashes over-densities along the line of sight in redshift space. This causes a boost in the amplitude of the redshift space correlation function on linear scales, as compared to the real space correlation function. This is characterised by the parameter $\beta = \Omega_m^{0.6}/b$ (Kaiser 1987), as defined in Chapter 3. In order to measure properly any effect of Λ, these two contributions have to be accurately accounted for in any model of the clustering. Models for the redshift space distortions have appeared many times in the literature (e.g. Bean et al. 1983, Ballinger, Peacock & Heavens 1996, Matsubara & Suto 1996, Popowski et al. 1998, Ratcliffe et al. 1998c), and we describe some of these models in Section 7.4.

Either galaxy or QSO surveys can, in theory, be used to detect the distortion introduced into the clustering pattern through incorrect assumptions for the cosmology. However, wide angle galaxy surveys, including the 2dF Galaxy Redshift Survey, essentially probe the Universe at $z = 0$ where assuming the incorrect cosmology would make little difference to the measured clustering pattern. Clustering from pencil beam surveys, such as the CNOC survey (Yee et al. 2000), cannot be measured over a wide range of scales. QSO surveys sample space more sparsely than current galaxy surveys but probe clustering out to high redshifts over a wide area, potentially allowing the distortions in the clustering to be detected.

The aim of this Chapter is to use the mock catalogues, constructed in Chapter 5, to test how well Ω_m, Ω_Λ and β will be constrained from the 2dF QSO survey. At this

stage, we only consider flat cosmologies with $\Omega_m + \Omega_\Lambda = 1$, as expected from current CMB experiments (Balbi et al. 2000, de Bernardis et al. 2000) and inflationary theory. The method can, however, be extended to allow for open or closed cosmologies. We also test whether tighter constraints on the cosmology can be found if the results from $\xi(\sigma, \pi)$ are combined with results from the evolution of the QSO-mass bias.

The outline of this Chapter is as follows. In Section 7.3, we discuss how the correlation function parallel and perpendicular to the line of sight, $\xi(\sigma, \pi)$, is measured from either the simulation or the 2dF QSO survey and in Section 7.4 we discuss how we model $\xi(\sigma, \pi)$. In Section 7.5, we outline our method for detecting the simulation cosmology (this is the method that will be applied to the 2dF QSO survey once it is completed to determine the underlying cosmology of the Universe) and in Section 7.6 we present our results. In Section 7.7, we consider whether other constraints can be usefully combined with the results from $\xi(\sigma, \pi)$ and finally in Section 7.8 we draw our conclusions. First, we define some of the terminology used in this Chapter.

7.2 Definitions

There are many terms used in this analysis that can be easily confused. We define their meaning here and stick to these conventions throughout.

- Underlying cosmology - this is the (unknown) cosmology of the Universe
- The simulation cosmology - the known cosmology (Ω_m=0.3, Ω_Λ=0.7) of the *Hubble Volume* Simulation
- Assumed cosmology - the cosmology used when measuring the two-point correlation function and $\xi(\sigma, \pi)$ from the 2dF QSO survey or the *Hubble Volume* simulation. Models of $\xi(\sigma, \pi)$, which are discussed later, also have to be calculated in the same assumed cosmology as the data. We consider two possibilities for the assumed cosmology, Ω_m=1, Ω_Λ=0 or Ω_m=0.3, Ω_Λ=0.7, when fitting the models to the simulation to show that the results are not sensitive to this choice.
- Test cosmology - the cosmology used to generate the model predictions for $\xi(\sigma, \pi)$ which are then translated into the assumed cosmology, discussed in Section 7.5. When the test cosmology matches the underlying (or simulation) cosmology, the distortions introduced into the clustering pattern should be the same in the model

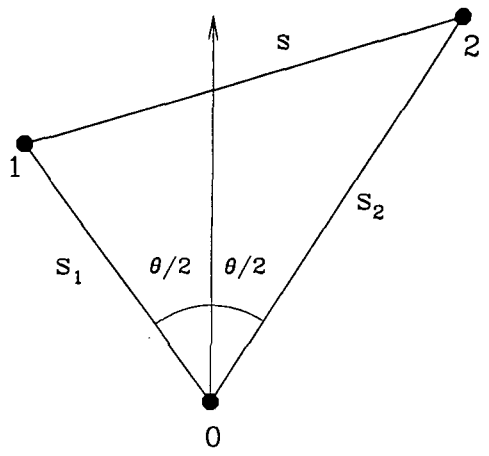

Figure 7.1: 1 and 2 are the positions of two QSOs in redshift space. s_1 and s_2 are the distances to the two QSOs measured from the observer at 0.

and in the data. The model should then provide a good fit to the data. If the test cosmology is incorrect then the model should not fit the data. This potentially allows cosmology to be constrained.

7.3 Measuring $\xi(\sigma, \pi)$

There are many different ways in which the clustering perpendicular (σ) and parallel (π) to the line of sight can be defined. Perhaps the simplest way is to define $\pi = |s_2 - s_1|$ and $\sigma = (s_1 + s_2)\theta/2$, where s_1 and s_2* are the distances to two QSOs and θ is the angle between them, measured from the position of an observer. Figure 7.1 shows the definitions of s_1, s_2 and θ. A more complicated approach is to work out their radial separation along the vector that bisects the angle between the two QSOs and to use Pythagoras' theorem to calculate σ, assuming that the Universe in any cosmology is locally flat. However, in practice it makes little difference which method of estimating π and σ is used as the clustering signal becomes very weak on scales larger than around $40h^{-1}$Mpc in either direction and differences between the two methods only become important when the angular or radial separation becomes large.

*We adopt the convention that s refers to apparent distances in redshift space and r refers to distances in real space through out this Chapter.

$\xi(\sigma,\pi)$ is then estimated in much the same way as the two-point correlation function, see Chapter 5. A catalogue of unclustered points, that have the same radial selection function and angular mask as the data but which contains many more points than the data catalogue, is used to estimate the effective volume of each bin. The DD, DR and RR counts, defined in Chapter 5, in each π and σ bin are found and the Hamilton estimator (Hamilton 1993), see also Chapter 5, is used to find $\xi(\sigma,\pi)$. Due to the sparsity of QSOs in the 2dF QSO survey, we use bins of $\delta\log(\pi/h^{-1}\mathrm{Mpc}) = \delta\log(\sigma/h^{-1}\mathrm{Mpc}) = 0.2$. As all the signal in the correlation function occurs on scales less than $s \lesssim 40h^{-1}\mathrm{Mpc}$, we adopt Poisson errors (again see Chapter 5, equation 5.8).

7.4 Modeling $\xi(\sigma,\pi)$

The aim of this Chapter is to test whether the cosmological parameters, Ω_m, Ω_Λ and also β will be constrained from the 2dF QSO Survey. β has been measured from galaxy redshift surveys by comparing the zeroth and second order moments of the correlation function (Ratcliffe et al. 1998c). However, the effects of cosmology add an extra distortion to the clustering pattern of QSOs. This alters the moments of the clustering too, such that the effects of redshift space and cosmological distortions would be very difficult to disentangle.

Instead, we test if cosmology and β will be constrained by comparing $\xi(\sigma,\pi)$ measured from the mock catalogues, created in Chapter 5, to models of $\xi(\sigma,\pi)$. The method is described in more detail in Section 7.5. The idea is that when the values of Ω_m, Ω_Λ and β used to calculate the models are the same as the underlying values, the model will match the mock catalogues, allowing the cosmology to be determined.

There are several assumptions that go into the models. One of the assumptions is that we are comparing the models to the data on linear scales so the effects of non-linear clustering can be ignored in the models. A second assumption we make is that the bias factor, b, is independent of scale. This is the case, by design, in the mock catalogues but this may only be an approximation on scales of $\sim 4h^{-1}\mathrm{Mpc}$ in the 2dF QSO survey.

The literature contains many examples of models for $\xi(\sigma,\pi)$ (e.g. Matsubara & Suto 1996, Popowski et al. 1998, Ratcliffe et al. 1998c) and the power spectrum, measured parallel and perpendicular to the line of sight, $P(k_{||},k_\perp)$, (Ballinger, Peacock & Heavens 1996). We consider various aspects of these models here. The models must account for the effects of cosmology and the redshift space distortions. We describe the method for

including the cosmology below and describe two models for the effects of redshift space distortions in Sections 7.4.2 and 7.4.3.

7.4.1 Effect of Λ on Clustering Anisotropy

If the assumed cosmology is different from the underlying cosmology of the Universe (or here the simulation), distortions, different from those caused by peculiar velocities, are introduced into the clustering pattern. This is because the radial and perpendicular directions are affected by cosmology in different ways. Ballinger, Peacock & Heavens (1996) outline how the cosmology scales the power spectrum split into its parallel and perpendicular components. Cosmology similarly affects the correlation function in the linear regime according to:

$$\xi(\sigma,\pi) = \xi(\sigma',\pi'), \tag{7.1}$$

i.e. the correlation function in the assumed cosmology is the same as the correlation function in the test cosmology provided σ and π are scaled according to

$$\sigma' = \frac{\sigma}{f_\perp} = \frac{\sigma B_\mathrm{a}}{B_\mathrm{t}} \tag{7.2}$$

and

$$\pi' = \frac{\pi}{f_\parallel} = \frac{\pi A_\mathrm{a}}{A_\mathrm{t}}, \tag{7.3}$$

where the subscripts a and t refer to the assumed cosmology and the test cosmology respectively. For a flat Universe, A and B are defined as

$$A = \frac{c}{H_0}\frac{1}{\sqrt{\Omega_\Lambda + \Omega_\mathrm{m}(1+z)^3}} \tag{7.4}$$

and

$$B = \frac{c}{H_0}\int_0^z \frac{dz'}{\sqrt{\Omega_\Lambda + \Omega_\mathrm{m}(1+z')^3}}. \tag{7.5}$$

These can be calculated for open universes too — see Ballinger, Peacock & Heavens (1996) for the full definition — however we only consider flat universes. Both A and B are calculated at the median redshift of the survey, which in this case is $z = 1.4$. By comparing the models to the simulation, we find that this is an adequate approximation to make, see Figures 7.3, 7.5 and 7.6.

7.4.2 Redshift Space Distortion Model 1

Following Peebles (1980) and Ratcliffe et al. (1998c) we define the correlation function, parallel and perpendicular to the line of sight by

$$1 + \xi(\sigma,\pi) = \int [1+\xi(r)]g(\mathbf{r},\mathbf{w})\mathrm{d}w^3, \tag{7.6}$$

where $\xi(r)$ is the real space correlation function, free from the effects of redshift space distortions. $\mathbf{w} = \mathbf{v_i} - \mathbf{v_j}$ where $\mathbf{v}$ is the peculiar velocity of a galaxy after subtracting off the Hubble flow and $\mathbf{r} = \mathbf{r_i} - \mathbf{r_j}$. $g(\mathbf{r}, \mathbf{w})$ is the distribution function of $\mathbf{w}$ for galaxies separated by $\mathbf{r}$. Here $r^2 = \sigma^2 + r_z^2$, where $r_z = \pi - w_z/H$ and w_z is the component of $\mathbf{w}$ parallel to the line of sight, denoted by z for simplicity.

If it can be assumed that $g(\mathbf{r}, \mathbf{w})$ is a slowly varying function of $\mathbf{r}$ then $g(\mathbf{r}, \mathbf{w}) = g(\mathbf{w})$. Then equation 7.6 can be simplified and equation 7.6 becomes

$$1 + \xi(\sigma, \pi) = \int_{-\infty}^{\infty} [1 + \xi(r)] f(w_z) \mathrm{d}w_z. \tag{7.7}$$

where

$$f(w_z) = \int \mathrm{d}w_x \int \mathrm{d}w_y g(\mathbf{w}) \tag{7.8}$$

A simple streaming model for the bulk motions of galaxies can be included by writing $g(\mathbf{r}, \mathbf{w}) = g[\mathbf{w} - \hat{\mathbf{r}} v(r)]$ where $v(r)$ is the mean relative radial velocity of galaxies separated by $\mathbf{r}$. If the approximation for $g(\mathbf{r}, \mathbf{w})$ in equation 7.8 is made again, equation 7.6 becomes

$$1 + \xi(\sigma, \pi) = \int_{-\infty}^{\infty} [1 + \xi(r)] f[w_z - v(r_z)] dw_z. \tag{7.9}$$

Models for the effects of small scale peculiar velocities and the bulk motions of galaxies or QSOs are required if $\xi(\sigma, \pi)$ is to be accurately described.

Fingers of God

The distribution function of the small scale peculiar velocities has been previously modeled as an exponential or a Gaussian distribution. Ratcliffe et al. (1998c) found that a substantially better fit to peculiar velocities in N-body simulations was found if an exponential model was used and that is what we adopt here:

$$f(w_z) = \frac{1}{\sqrt{2} < w_z^2 >^{1/2}} \exp\left\{ -\sqrt{2} \frac{|w_z|}{< w_z^2 >^{1/2}} \right\}, \tag{7.10}$$

where $< w_z^2 >^{1/2}$ is the rms line of sight pairwise velocity dispersion. The value of the rms line of sight pairwise velocity dispersion is unknown for the 2dF QSOs and is therefore left as a free parameter in the model.

Bulk Flows

The model that we use for the bulk motions depends on the cosmology of the Universe, the clustering of QSOs and the QSO bias. Following Hale-Sutton (1990) we set

$$v(r_z) = -\frac{2}{3 - \gamma} \Omega_{\mathrm{m}}^{0.6} H(z) r_z \left\{ \frac{\xi(r)}{b^2 + \xi(r)} \right\}. \tag{7.11}$$

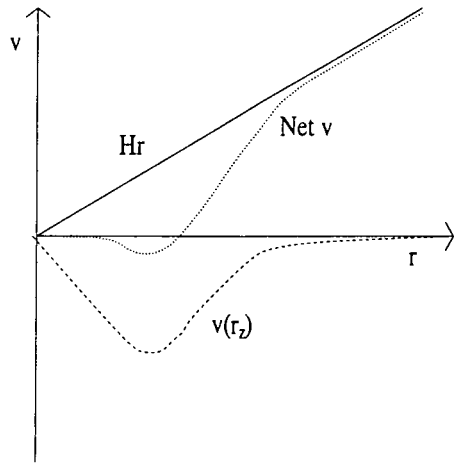

Figure 7.2: Diagram to show how objects move according to the model for the bulk flow. The straight line (solid) shows the Hubble Law, Hr. The lowest line (dashed) shows the mean radial velocity, as given by equation 7.11. The middle line (dotted) shows the net motion of objects.

r_z is defined above, β and Ω_m are free parameters and the bias, b, is calculated using $b = \Omega_m^{0.6}/\beta$ (Kaiser 1987). H_o=70 km s^{-1} Mpc^{-1} and the evolution of Hubble's constant with redshift and cosmology is given in Appendix A. We self-consistently determine the real space correlation function from the redshift space correlation function. This is discussed in Section 7.5. The real space correlation function is not approximated by a power law but a value of γ is still required in equation 7.11. A value of γ=1.7 best fits the redshift space correlation function over the range $4 < s < 40h^{-1}$Mpc (see Chapter 5 and Figure 7.8) so this is the value used in equation 7.11.

On small scales, $\xi(r) \gg 1$, so equation 7.11 simplifies to

$$v(r_z) = -\frac{2}{3-\gamma}\Omega_m^{0.6} H r_z. \tag{7.12}$$

The term $2\Omega_m^{0.6}/(3-\gamma)$ is of order unity so the Hubble flow is to some extent cancelled by the bulk motions, as might be expected in a stable cluster. On large scales, $\xi(r) \ll 1$ so equation 7.11 tends to zero and on average, objects are moving away from us with the Hubble Flow. This is summarised in Figure 7.2. The lower line (dashed) shows the mean inward radial velocity, as given by equation 7.11. The upper line (solid) shows the Hubble Law, Hr. The middle line (dotted) shows the net motion of objects.

7.4.3 Redshift Space Distortion Model 2

An alternative method for modeling the effects of infall on $\xi(\sigma,\pi)$ is outlined in Matsubara & Suto (1996). Essentially, they generalise the formula of Hamilton (1992) to define a relation between the redshift space correlation function at $z \neq 0$ and the real space correlation function. By expanding the redshift space correlation function as a series of Legendre polynomials, Matsubara & Suto (1996) show that

$$\xi(\sigma,\pi) = \xi(r)\left\{1 + \frac{2(1-\gamma\mu^2)}{3-\gamma}\beta(z) + \frac{\gamma(\gamma+2)\mu^4 - 6\gamma\mu^2 + 3}{(3-\gamma)(5-\gamma)}\beta(z)^2\right\}, \qquad (7.13)$$

where $r^2 = \sigma^2 + \pi^2$ and $\mu = \pi/r$ and γ is the same as described in Model 1. This model includes the effects of infall but not small scale peculiar velocities. These are included by convolving the redshift space correlation function with an exponential model for small scale peculiar velocities along the π direction.

7.4.4 Comparison with the Hubble Volume Simulation

Do the two models described above reproduce $\xi(\sigma,\pi)$ measured from the mock QSO catalogues? Figure 7.3 shows $\xi(\sigma,\pi)$ measured from the mock catalogues. We show the fractional errors on $\xi(\sigma,\pi)$ in Figure 7.4 — the lighter the shading, the smaller the errors. The areas of the $\xi(\sigma,\pi)$ diagram where the fractional errors are the smallest are the areas where most of the differentiation between different models of $\xi(\sigma,\pi)$ can be made. Figure 7.5 shows $\xi(\sigma,\pi)$ calculated using Model 1 and Figure 7.6 shows $\xi(\sigma,\pi)$ calculated using Model 2. Three different cosmologies have been assumed for the simulation and the models, $\Omega_{\rm m}$=1, Ω_Λ=0 (shown in panels a, d and g), $\Omega_{\rm m}$=0.3, Ω_Λ=0.7 (panels b, e and h), which is also the simulation cosmology, and $\Omega_{\rm m} = 0$, $\Omega_\Lambda = 1$ (panels c, f and i). The bold, solid lines show ξ=0.1, the dot dashed bold lines show ξ=1. The solid lines increase from ξ=0.1 in steps of 0.1, the dashed lines decrease from $\xi = 0.1$ in steps of 0.01 and the dot-dashed lines increase from ξ=1 in steps of 1. The values of the velocity dispersion ($< w_z^2 >$=400 km s^{-1}) and β (0.36) in the models are the values measured from the Hubble Volume simulation at the average redshift and the real space correlation function in the models is taken directly from the *Hubble Volume* simulation. The models should fit the simulation in each assumed cosmology as closely as possible.

The agreement between the models and the measurement of $\xi(\sigma,\pi)$ from the Hubble Volume simulation is fairly good. Table 7.1 shows the results of χ^2 fitting the model with

Model	Assumed Cosmology	χ^2
Model 1	Ω_m=1 Ω_Λ=0	23.6
	Ω_m=0.3 Ω_Λ=0.7	24.7
	Ω_m=0 Ω_Λ=1	24.3
Model 2	Ω_m=1 Ω_Λ=0	24.1
	Ω_m=0.3 Ω_Λ=0.7	25.1
	Ω_m=0 Ω_Λ=1	24.5

Table 7.1: We show the goodness of fit of the model $\xi(\sigma,\pi)$ to the simulation $\xi(\sigma,\pi)$ with the assumed cosmologies consistent. Model 1 provides a slightly better fit to the simulation than Model 2, although the differences are fairly small. The fits were made to 25 bins with two degrees of freedom, the test cosmology and β.

the assumed cosmology matching that used to measure the clustering from the *Hubble Volume* simulation. Model 1 provides a slightly better fit to the simulation than Model 2, although the differences between the two models are fairly small. All of the models match the simulation to within $\sim 1\sigma$. For simplicity we adopt Model 1 throughout.

The measurement of $\xi(\sigma,\pi)$ from the Hubble Volume simulation is quite noisy. However, particularly if the bold, solid line ($\xi = 0.1$) is considered, the effects of the different assumed cosmologies can be seen. This line in panel a) is far more squashed in the π direction than in panel b) and elongation is seen in panel c) as compared to panel b). This is seen in the models too, Figures 7.5 (Model 1) and 7.6 (Model 2). The solid line in panels d) and g) appears more squashed in the π direction than the lines in panels e) and h) and the lines in panels f) and i) appear more elongated in the π direction than the lines in panels e) and h).

7.5 Determining the Underlying (Simulation) Cosmology

The underlying cosmology of the Universe is not known. When a correlation function is measured from a redshift survey a cosmology has to be assumed in order for the separations of galaxies or QSOs to be calculated. Out to redshifts of $z \sim 0.1 - 0.2$ assuming the wrong cosmology leads to separations being estimated incorrectly by around 10% for flat cosmologies. However, if clustering is measured over a wider range of redshifts, the effects of assuming the wrong cosmology become large, as can be seen in Figures 7.3, 7.5 and 7.6. This potentially allows the cosmology to be detected from the clustering pattern

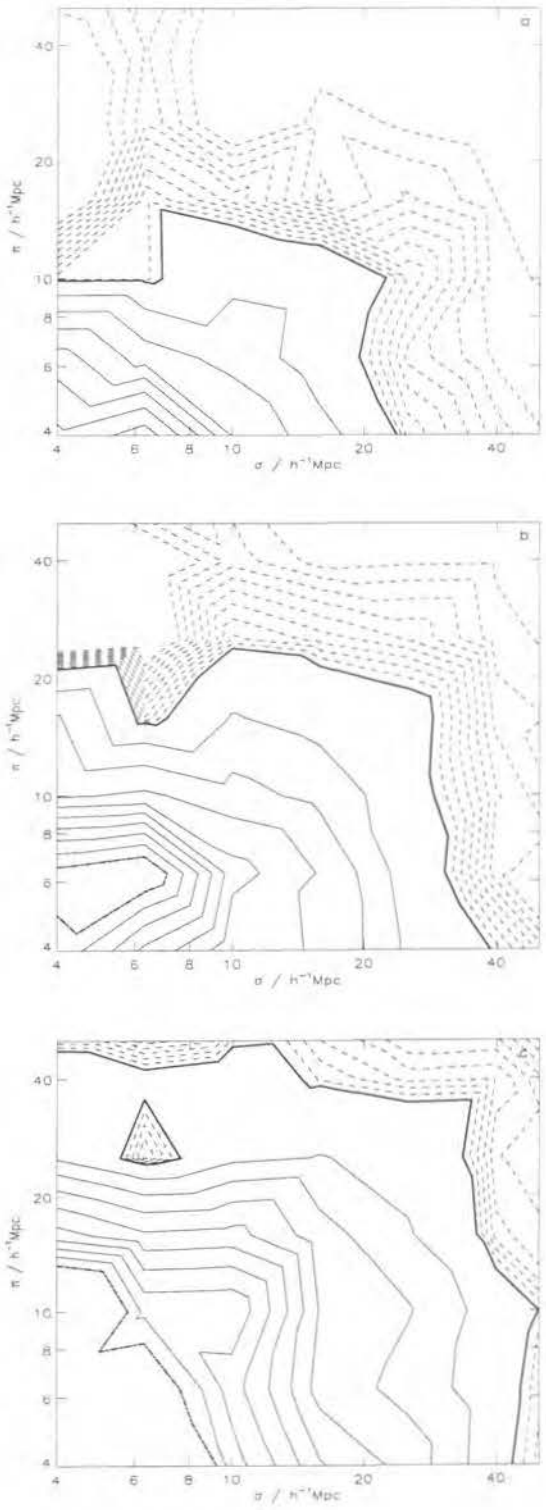

Figure 7.3: $\xi(\sigma, \pi)$ estimated from the *Hubble Volume* simulation. Each strip contains 12,500 mock QSOs and the average over the three strips is shown. The simulation cosmology is Ω_m=0.3, Ω_Λ=0.7 and the assumed cosmology is a) Ω_m=1, Ω_Λ=0 b) Ω_m=0.3, Ω_Λ=0.7 and c) $\Omega_m = 0$, $\Omega_\Lambda = 1$. The solid, bold contour shows $\xi = 0.1$ and the solid, dot-dashed contour shows $\xi = 1$. See the text for the value of the other lines.

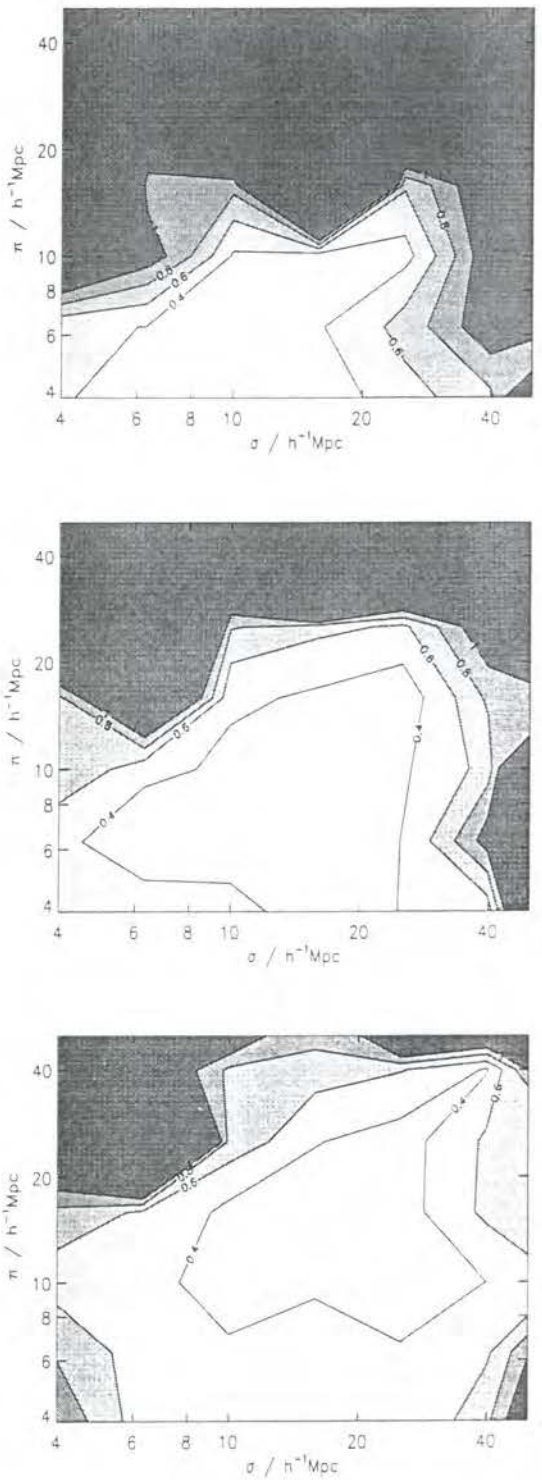

Figure 7.4: The fractional errors on $\xi(\sigma, \pi)$ measured from the *Hubble Volume* simulation assuming Ω_m=1, Ω_Λ=0 (top) Ω_m=0.3, Ω_Λ=0.7 (middle) and $\Omega_m = 0$, $\Omega_\Lambda = 1$ (bottom). The white area shows where $\xi(\sigma, \pi)$ is measured to better than 40%.

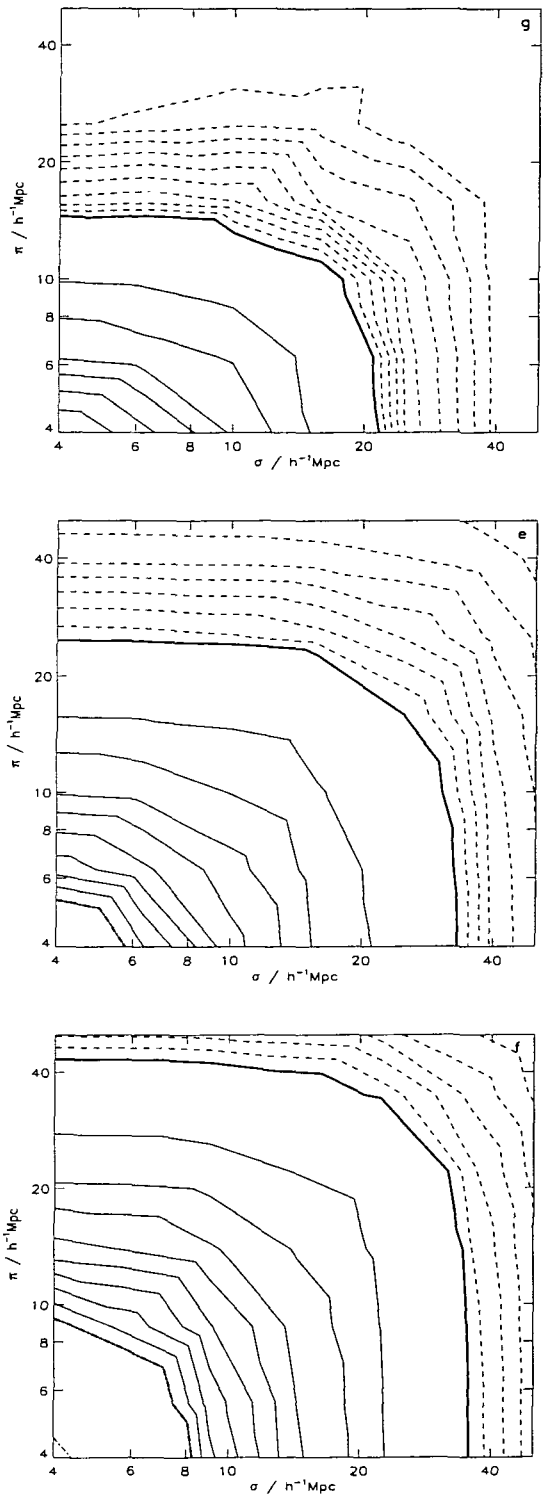

Figure 7.5: $\xi(\sigma, \pi)$ estimated using Model 1. The value of the velocity dispersion, β and the real space correlation function are taken from the *Hubble Volume* simulation. The test cosmology is Ω_m=0.3, Ω_Λ=0.7 and the assumed cosmology is d) Ω_m=1, Ω_Λ=0 e) Ω_m=0.3, Ω_Λ=0.7 and f) $\Omega_m = 0$, $\Omega_\Lambda = 1$. The values of the ξ contours are described in the text.

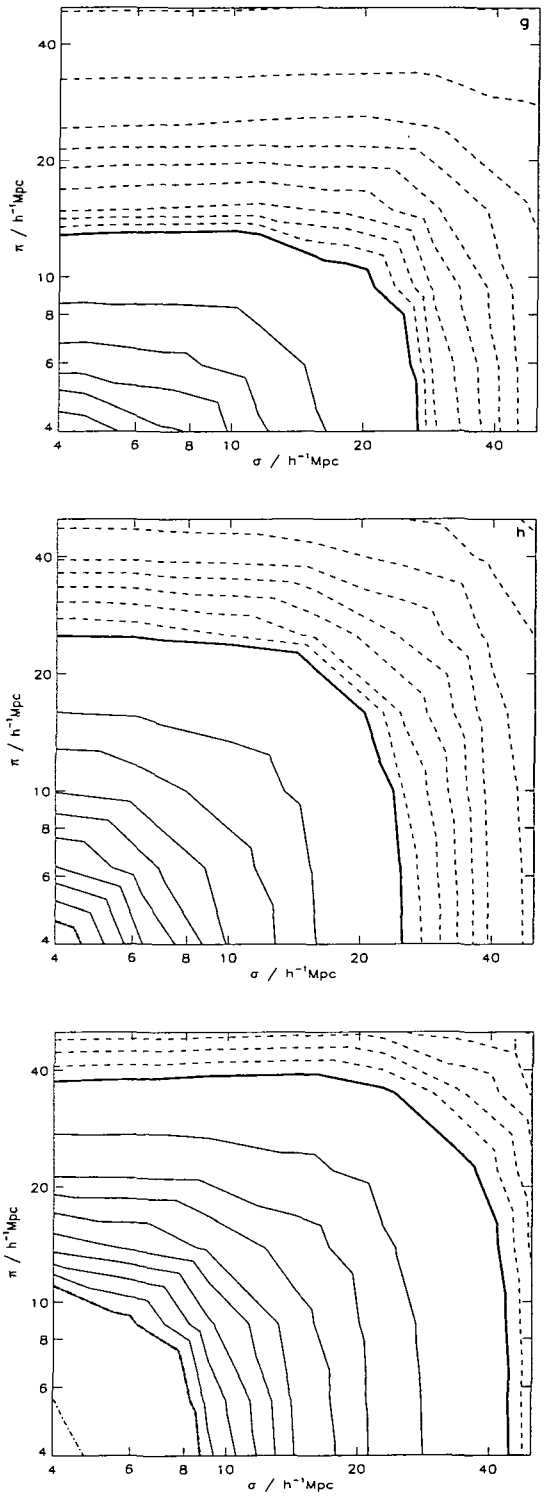

Figure 7.6: $\xi(\sigma, \pi)$ estimated using Model 2. The value of the velocity dispersion, β and the real space correlation function are taken from the *Hubble Volume* simulation. The test cosmology is again Ω_m=0.3, Ω_Λ=0.7 and the assumed cosmology is g) Ω_m=1, Ω_Λ=0 h) Ω_m=0.3, Ω_Λ=0.7 and i) $\Omega_m = 0$, $\Omega_\Lambda = 1$. The values of the ξ contours are described in the text.

of QSOs.

By comparing a range of models with different test cosmologies to $\xi(\sigma,\pi)$ measured from the 2dF QSO survey the underlying cosmology of the Universe can, in principle, be detected. This is because when the test cosmology is the same as the underlying cosmology in the Universe, the same distortions due to cosmology will be introduced into the clustering pattern of the model and the clustering pattern of the 2dF QSO Survey, as the values of the assumed cosmology are set to be the same. The model $\xi(\sigma,\pi)$ will then give a good fit to $\xi(\sigma,\pi)$ measured from the 2dF QSO survey. This is providing, of course, that the redshift space distortions are properly accounted for. However, the values of β and the small scale velocities are not known, neither is the real space correlation function. Therefore, a grid of models with varying cosmology, β, small scale velocities and real space correlation function is required for comparisons to the data.

However, as we compare the models and the simulation on scales greater than $4h^{-1}$Mpc in both the π and σ direction, we find that the modeling of $\xi(\sigma,\pi)$ is fairly insensitive to the value of the small scale velocity dispersion. This is shown in Figure 7.7. Here we show two different sets of models, the top plot shows models with test cosmology $\Omega_{\rm m}$=1, Ω_Λ=0, β=0.36 and an assumed cosmology $\Omega_{\rm m}$=0.3, Ω_Λ=0.7. The solid lines have $< w_z^2 >=$ 50 km s^{-1} and the dashed lines have $< w_z^2 >=$ 1000 km s^{-1}. The lower plot shows models with test cosmology $\Omega_{\rm m}$=0.3, Ω_Λ=0.7, β=0.36 and an assumed cosmology $\Omega_{\rm m}$=0.3, Ω_Λ=0.7. Again the solid lines have $< w_z^2 >=$ 50 km s^{-1} and the dashed lines have $< w_z^2 >=$ 1000 km s^{-1}. Considering the wide range of velocities covered by these two models, there is little difference between them apart from on the smallest σ scales where the errors on $\xi(\sigma,\pi)$ from the mock catalogues are large anyway, see Figure 7.4.

$< w_z^2 >$ will not be known from the 2dF QSO Survey. To quantify this effect, in the case where the test cosmology is $\Omega_{\rm m}$=0.3, Ω_Λ=0.7, we have χ^2 fitted the models to the simulation. For both choices of $< w_z^2 >$ shown in Figure 7.7, the value of χ^2 is increased by less than 2 compared to the χ^2 value for $< w_z^2 >$=400 km s^{-1} shown in Table 7.1. An increase of less than 2 in the obtained value of χ^2, corresponds to the model still matching the simulation to within 1σ. This is mainly because it is only on small $\sigma \lesssim 8h^{-1}$Mpc scales that the peculiar velocities have a significant effect on the shape of $\xi(\sigma,\pi)$. From Figure 7.4, we see that most of the differentiation between different models of $\xi(\sigma,\pi)$ occurs on larger scales. We therefore fix $< w_z^2 >=$ 400 km s^{-1} in the models.

As the effects of small scale peculiar velocities have little impact on models of $\xi(\sigma,\pi)$ on scales larger than $4h^{-1}$Mpc, they should also have little effect on the correlation

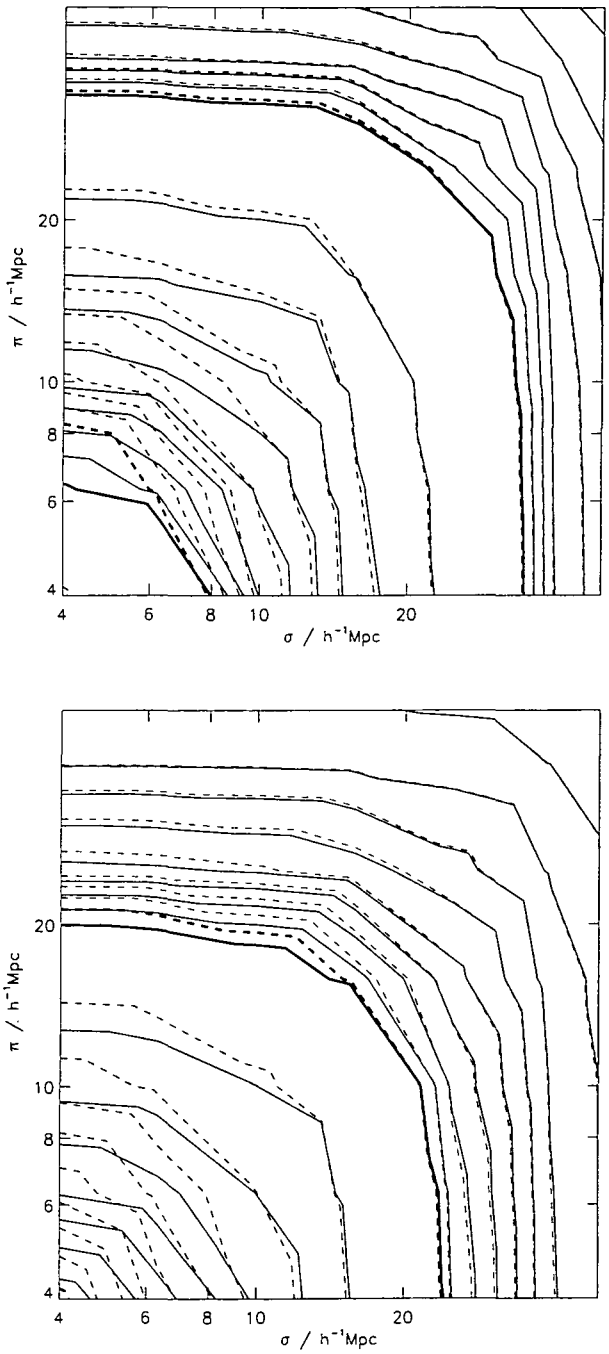

Figure 7.7: Two sets of models with different values of the small scale peculiar velocities. The top plot shows models with test cosmology Ω_m=1, Ω_Λ=0., β=0.36 and an assumed cosmology Ω_m=0.3, Ω_Λ=0.7. The solid lines have $< w_z^2 >=$ 50 km s^{-1} and the dashed lines have $< w_z^2 >=$ 1000 km s^{-1}. The lower plot shows models with test cosmology Ω_m=0.3, Ω_Λ=0.7, β=0.36 and an assumed cosmology Ω_m=0.3, Ω_Λ=0.7. Again the solid lines have $< w_z^2 >=$ 50 km s^{-1} and the dashed lines have $< w_z^2 >=$ 1000 km s^{-1}. Apart from on the smallest scales, there is little difference between the two models considering the large difference between the two velocities and the error contours shown in Figure 7.4.

function $\xi(s)$. We have tested this on the *Hubble Volume* simulation by comparing the real and redshift space correlation functions of the mass. We find that on scales greater than $\sim 3h^{-1}$Mpc the small scale peculiar velocities have negligible effect on the shape of the redshift space mass correlation function. We measure $< w_z^2 > \sim 400$ km s^{-1} for the biased particles at $z = 1.4$ from the *Hubble Volume* simulation, shown in Figure 5.3. The value of the velocity dispersion at $z = 0$ is ~ 600 km s^{-1}. However, galaxies have a lower value for the pairwise velocity dispersion of ~ 400 km s^{-1} at $z = 0$ (Ratcliffe et al. 1998c, see also Chapter 3). Therefore, if the peculiar velocities of QSOs are similar to those of galaxies and if the peculiar velocities of QSOs decreases with redshift, similar to what is seen in the mock catalogues, then it is possible that $< w_z^2 >= 400$ km s^{-1} at $z = 1.4$ will be an upper limit.

It should therefore be possible to find the real space correlation function self-consistently from the redshift space correlation function for each test value of β. For each test value of β, the real space correlation function can be found via the formula of Kaiser (1987)

$$\xi(r) = \frac{\xi(s)}{\left\{1 + \frac{2}{5}\beta + \frac{1}{5}\beta^2\right\}}, \tag{7.14}$$

where r and s denote the correlation function in real and redshift space respectively and $\beta = \Omega_{\rm m}^{0.6}/b$. This is measured in the assumed cosmology but, to create the model $\xi(\sigma, \pi)$, we need the real space correlation function measured in the test cosmology. The value of $\xi(r)$ in the test cosmology is the same as $\xi(r')$ in the assumed cosmology with

$$r^2 = \sigma^2 + r_z^2 \tag{7.15}$$

and

$$r'^2 = (\sigma/f_\perp)^2 + (r_z/f_{||})^2. \tag{7.16}$$

The terms $f_\perp$ and $f_{||}$ are defined in equation 7.2 and 7.3. The accuracy of converting the two-point correlation function from one cosmology into another is demonstrated in Figure 7.8. The symbols show the redshift space correlation function measured from the *Hubble Volume* with $\Omega_{\rm m}$=1 (triangles) and $\Omega_{\rm m}$=0.3, Ω_Λ=0.7 (circles). The line shows the correlation function measured in the $\Omega_{\rm m}$=1 cosmology but translated into the $\Omega_{\rm m}$=0.3, Ω_Λ=0.7 cosmology. There is good agreement between the two over a wide range of scales even though the correlation function is measured with fairly coarse binning. This is a further consistency check that measuring $f_\perp$ and $f_{||}$ at the average redshift of the Survey is an adequate approximation to make. On the smallest, non-linear scales the translation will not scale the correlation function correctly as the effects of the small scale peculiar

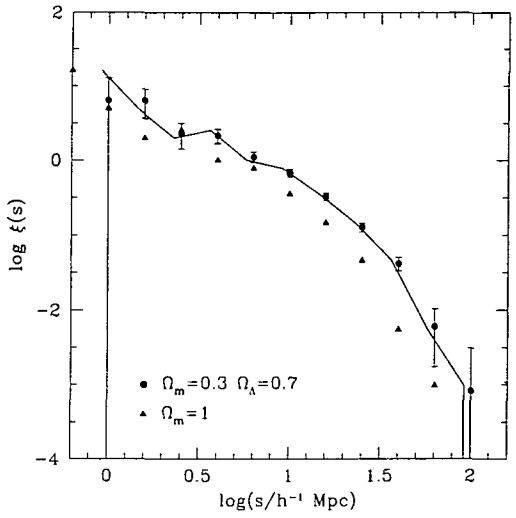

Figure 7.8: s

The correlation function of the mock catalogues estimated assuming Ω_m=0.3, Ω_Λ=0.7 (circles) and Ω_m=1, Ω_Λ=0 (triangles). The line shows the correlation function measured assuming Ω_m=1 corrected to a cosmology with Ω_m=0.3, Ω_Λ=0.7 with the relation between r and r' given in equation 7.16. This line agrees well with the circles over the range of scale $3< r <50h^{-1}$Mpc, showing that a correlation function measured in one cosmology can be translated into another.

velocities may be different in each cosmology. We will struggle to measure $\xi(\sigma,\pi)$ from the 2dF QSO Survey on these scales due to redshift errors (see Section 7.6.1 for more details).

The fitting procedure that we adopt to find which test cosmology matches the simulation cosmology is as follows:

1) Pick an assumed cosmology (here either Ω_m=1, Ω_Λ=0 or Ω_m=0.3, Ω_Λ=0.7)

2) Calculate $\xi(s)$ and $\xi(\sigma,\pi)$ from the data (the 2dF QSO survey or, in this case, the mock catalogues) using the assumed cosmology.

3) Pick a value for the test $\beta(\bar{z})$ at the average redshift of the survey.

4) Pick a value of the present day $\Omega_m(0)$ and calculate $\Omega_m(\bar{z})$ at $\bar{z}$=1.4 via

$\Omega_m(\bar{z}) = \Omega_m(0)(1+\bar{z})^3/[\Omega_m(0)(1+\bar{z})^3 + \Omega_\Lambda(0)]$.

5) Calculate the bias $b = \Omega_m^{0.6}(\bar{z})/\beta(\bar{z})$.

6) Translate the redshift space correlation function, $\xi(s)$, measured in the assumed cosmology to a real space correlation function, $\xi(r)$, in the test cosmology

using equation 7.16 and 7.14.

7) Generate the model $\xi(\sigma,\pi)$ using Model 1.

8) Translate the model $\xi(\sigma,\pi)$ from the test cosmology into the assumed cosmology using equations 7.3 and 7.2.

9) Calculate how well the model $\xi(\sigma,\pi)$ fits the data $\xi(\sigma,\pi)$ via the χ^2 statistic, using the Poisson errors from the data $\xi(\sigma,\pi)$ measured in the assumed cosmology.

10) Go back to 3) using a different test cosmology and $\beta(\bar{z})$.

When the parameters β and $\Omega_{\rm m}$ match those of the underlying cosmology (or in this case the simulation cosmology), the value of χ^2 should be minimised. We fit the model $\xi(\sigma,\pi)$ to the *Hubble Volume* $\xi(\sigma,\pi)$ over the range of scales $4< \sigma,\pi <40h^{-1}$ Mpc. This is to ensure that any non-linear effects are small and that the errors on $\xi(\sigma,\pi)$ do not dominate the actual value of $\xi(\sigma,\pi)$.

7.5.1 Calculating the 1σ Error Contours

The error contours are found as follows. The number of degrees of freedom, ν, is the number of bins in which the model $\xi(\sigma,\pi)$ is fitted to the data $\xi(\sigma,\pi)$ minus the number of free parameters (this is 2, $\Omega_{\rm m}$ and β). For this value of ν, the values of $\delta\chi^2$ that correspond to 1σ, 2σ etc. can be calculated. The minimum value of χ^2, found from fitting the grid of $\xi(\sigma,\pi)$ models to the simulation $\xi(\sigma,\pi)$, is close to the expected value, see Table 7.1. We then set the 1σ contour at $\chi^2_{\rm min}+\delta\chi^2(1\sigma)$ and similar for the 2 and 3σ contours.

We assume Poisson errors for $\xi(\sigma,\pi)$ measured from the *Hubble Volume*. As we only consider the clustering out to $s \lesssim 50h^{-1}$Mpc, these errors should be acceptable. If we were measuring the clustering out to larger scales than the errors would underestimate the true error (see Figure 5.6 in Chapter 5). Ideally, we would like many mock catalogues so that the dispersion over many realisations could be used for the errors. Currently, this is not possible.

7.6 Predicted Results for the 2dF QSO Survey

The 2dF QSO survey is not yet completed. Until the survey is finished, the measurement of $\xi(\sigma,\pi)$ is too noisy to determine the cosmology. Instead we predict at what level cosmology will be constrained from the completed 2dF QSO survey by considering the mock 2dF QSO catalogues constructed in Chapter 5.

The results of fitting the $\xi(\sigma,\pi)$ models to $\xi(\sigma,\pi)$ measured from the mock catalogues are given in Figure 7.9. The upper panel (a) shows the results when the assumed cosmology is Ω_m=1, Ω_Λ=0 and the lower panel (b) is for Ω_m=0.3, Ω_Λ=0.7. We show the results for two different assumed cosmologies to show they are not dependent on this choice. In both cases, the simulation cosmology, Ω_m=0.3, Ω_Λ=0.7 and $\beta = 0.36$ is contained within the 1σ contour. However, the results do not place a strong constraint on Ω_m as there is a degeneracy between Ω_m and $\beta(z)$. However, the value of $\beta(z)$ for QSOs at $z = 1.4$ should be fairly well constrained from the 2dF QSO Survey. This does at least place a joint constraint on cosmology and the QSO-mass bias.

The reason for this degeneracy is that the errors on $\xi(\sigma,\pi)$ from the mock catalogues are fairly large. Shown in Figure 7.10 are two different $\xi(\sigma,\pi)$ models. One has $\Omega_m = 0.3$, Ω_Λ=0.7 and a value of β=0.35 (dashed lines) and another model has Ω_m=1, Ω_Λ=0 and β=0.5 (solid lines). All the other parameters (the assumed cosmology is $\Omega_m = 0.3$, Ω_Λ=0.7, and the small scale peculiar velocities $< w_z^2 >$=400 km s^{-1} etc) are the same in each model. These models both provide a good fit to the mock catalogues from the *Hubble Volume* (see Figure 7.9) which is not surprising as the two models have similar shape and amplitude. Very small errors on $\xi(\sigma,\pi)$ are needed to detect the small differences between these models.

The degeneracy between Ω_m and β seen in Figure 7.9 does not lie in the direction that is perhaps intuitively expected. Consider the two assumed cosmologies adopted here, the two test cases shown in Figure 7.10 and the simulation cosmology Ω_m=0.3, Ω_Λ=0.7. If Ω_m=1, Ω_Λ=0 is assumed then $\xi(\sigma,\pi)$ from the simulation is squashed (seen in Figure 7.3(a)), as is the model $\xi(\sigma,\pi)$ with test cosmology Ω_m=0.3, Ω_Λ=0.7, due to the simulation and test cosmologies being different to the assumed cosmology. The test cosmology with Ω_m=1, Ω_Λ=0 is not compressed as the test cosmology matches the assumed cosmology. A higher value of β is therefore required to squash the model $\xi(\sigma,\pi)$ to match the simulation.

If Ω_m=0.3, Ω_Λ=0.7 is assumed, then no squashing due to cosmology occurs on the simulation $\xi(\sigma,\pi)$ (again seen in Figure 7.3(b)) or on the model $\xi(\sigma,\pi)$ with Ω_m=0.3, Ω_Λ=0.7 as the test cosmology as in this case the simulation and test cosmologies match the assumed cosmology. However, if Ω_m=1, Ω_Λ=0 is the test cosmology, then the model $\xi(\sigma,\pi)$ is elongated due to the different test and assumed cosmologies. Therefore, a larger value of β is again required to compensate for the elongation, allowing the model to match the simulation $\xi(\sigma,\pi)$. This is partly shown in Figure 7.7 as here we have

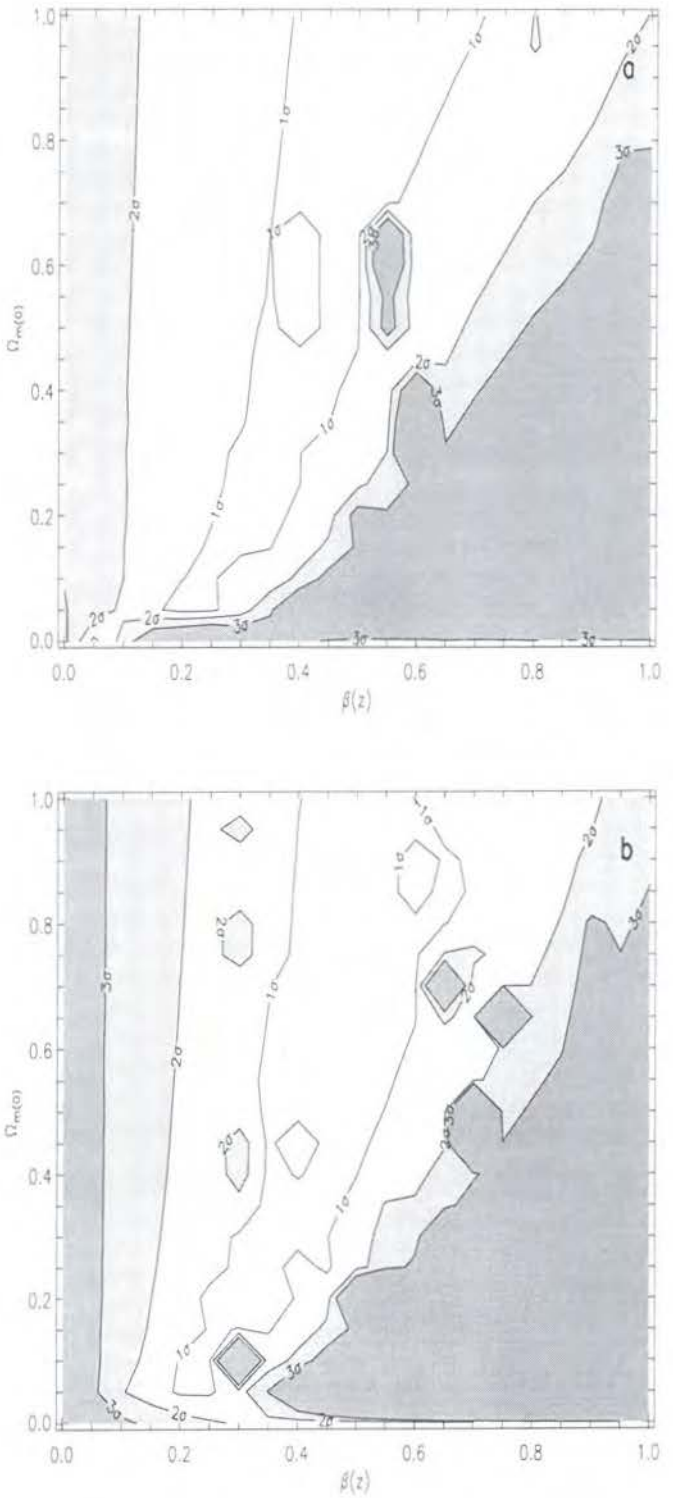

Figure 7.9: Fitting $\xi(\sigma, \pi)$ from the *Hubble Volume*. The top panel (a) has Ω_m=1 as the assumed cosmology and the lower panel (b) has Ω_m=0.3, Ω_Λ=0.7 as the assumed cosmology. In both cases the simulation cosmology, Ω_m=0.3, Ω_Λ=0.7 and β=0.36 is contained within the 1σ contour, although unfortunately little constraint on Ω_m is possible from the mock 2dF QSO Survey catalogues.

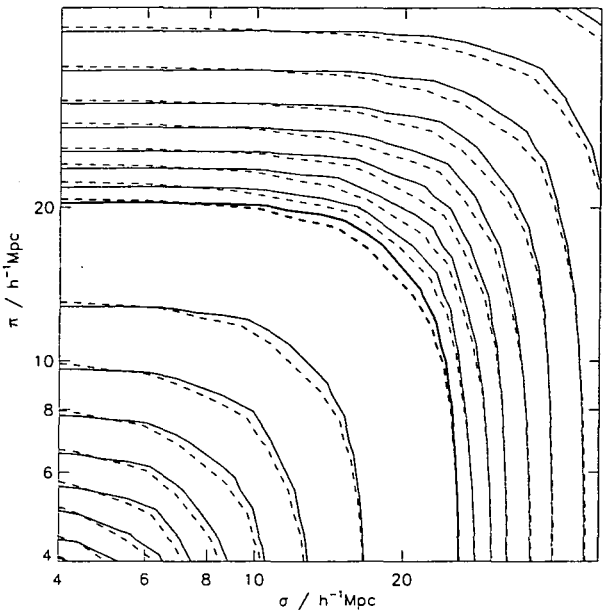

Figure 7.10: The dashed lines show a model with test cosmology Ω_m=0.3 and Ω_Λ=0.7 with a value of β=0.35. The solid lines show a model with test cosmology Ω_m=1 and β=0.5. The difference between the two models is small. The assumed cosmology in both cases is Ω_m=0.3 and Ω_Λ=0.7

shown models with assumed cosmology Ω_m=0.3, Ω_Λ=0.7, β=0.35 (although the small scale peculiar velocities are 50 and 1000 km s^{-1} rather than 400 km s^{-1}) but the top plot has test cosmology Ω_m=1, Ω_Λ=0 and the lower plot has test cosmology Ω_m=0.3, Ω_Λ=0.7. Elongation is clearly seen in the top plot as compared to the bottom plot, showing that a higher value of β is required for the model with test cosmology Ω_m=1, Ω_Λ=0 if the model is to match the simulation.

In Figure 7.11, we show an example of a model that is rejected with 1-2σ confidence and a model that is rejected with more than 3σ confidence. We compare these two models to a model with a test cosmology of Ω_m=0.3, Ω_Λ=0.7 and β=0.35 (dashed lines), which closely matches the simulation $\xi(\sigma, \pi)$. In both panels, the thick lines correspond to ξ=0.1 and 0.01 with ξ=0.1 appearing always to the left of the contour with ξ=0.01 (which lies just off the plot for the model with a test cosmology of Ω_m=0.3, Ω_Λ=0.7 and β=0.35). In the top panel, the solid lines show a model with a test cosmology of Ω_m=1, Ω_Λ=0 and β=0.2. This model is rejected at the $1-2\sigma$ level in Figure 7.9. On large scales the solid contours are more tightly packed than the dashed contours and on small scales, more elongation is seen in this model than in the model that closely matches the simulation. The lower panel shows a model with a test cosmology of Ω_m=0.1, Ω_Λ=0.9 and β=0.6

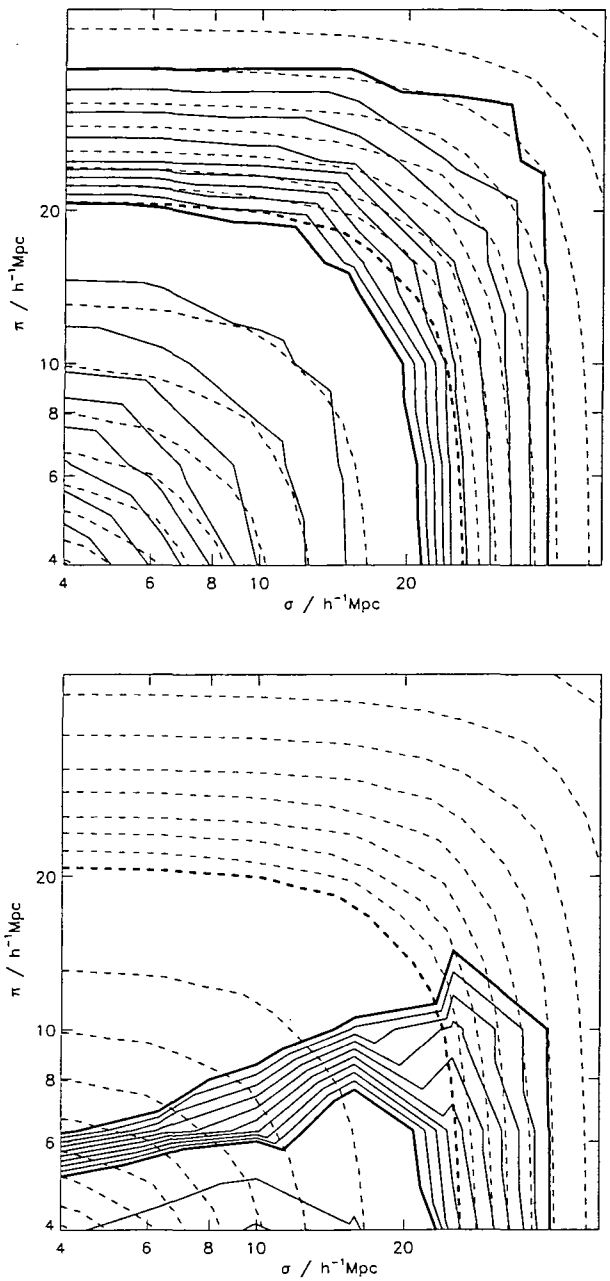

Figure 7.11: In both panels, the dashed lines show a model with test cosmology Ω_m=0.3 and Ω_Λ=0.7 with a value of β=0.35. This model closely matches the simulation $\xi(\sigma, \pi)$. The solid lines in the top panel show a model with test cosmology Ω_m=1 and β=0.2. This model differs from the simulation at the 1-2σ level. The lower panel shows a model with test cosmology Ω_m=0.1, Ω_Λ=0.9 and β=0.6 which clearly does not match the simulation. This model is rejected at more than 3σ. Thick contours show ξ=0.1 and 0.01 (ξ=0.01 lies just off the plot when the model is calculated with test cosmology Ω_m=0.3, Ω_Λ=0.7 and β=0.35). Ω_m=0.3 and Ω_Λ=0.7 is the assumed cosmology in all cases.

(solid lines). This model is rejected at more than 3σ in Figure 7.9 and is clearly seen to be discrepant from the model that closely matches the simulation. Ω_m=0.3, Ω_Λ=0.7 is the assumed cosmology in all of the above mentioned models.

7.6.1 Will things be any better from the completed 2dF QSO Survey?

The results for determining the cosmology from the mock catalogues suggest that the cosmological parameters, Ω_m and Ω_Λ, will not be well constrained from the 2dF QSO survey. However, the fits are only made to the mock 2dF QSO catalogues and there may be ways in which constraints can be improved when the 2dF QSO survey is completed.

Early results on the correlation function of the incomplete 2dF QSO survey suggest that the correlation function is higher on smaller scales than the mock catalogues (see Chapter 5, Figure 5.5). This is mainly because the biasing scheme was chosen to match the clustering of the 2dF QSO correlation function on scales around $10h^{-1}$Mpc rather than on smaller scales. The number of pairs found in the 2dF QSO Survey at small separations is therefore larger than the number found in the mock catalogues, which means that on the smallest scales, $s \sim 7h^{-1}$Mpc, the errors on $\xi(\sigma,\pi)$ from the 2dF QSO survey should be smaller. This may help to constrain cosmology slightly better. The fitting of $\xi(\sigma,\pi)$ models to the 2dF QSO Survey will not be possible on smaller scales than used here though. This is to because the QSO redshifts can only be measured to an accuracy of around 0.003 (Scott Croom, private communication). The inaccuracy in measuring the redshift arises from the fact that QSOs have broad emission lines so the redshift cannot be precisely measured. At the average redshift of the Survey, the error in the redshift corresponds to an accuracy of $\sim 4h^{-1}$Mpc in the Ω_m=0.3, Ω_Λ=0.7 cosmology adopted here. Measurements of $\xi(\sigma,\pi)$ are smoothed in the π direction on scales of $\sim 4h^{-1}$Mpc. The corresponding scale in the Ω_m=1 cosmology is $\sim 3h^{-1}$Mpc. Fitting the models to the data on scales less than $4h^{-1}$Mpc would, in any case, require us to account for the small scale peculiar velocities and non-linear effects more accurately.

The 2dF QSO Survey correlation function also seems to have more large scale power than the correlation function from the mock catalogues. This may mean that $\xi(\sigma,\pi)$ can be measured and compared to models out to larger scales. More bins may be used in the fitting which could lead to better discrimination between different cosmologies. Poisson errors would not be valid on these scales though so other error estimates would need to be used.

7.7 Further Constraints on Cosmology and β

As shown in Section 7.6, the cosmological parameters, Ω_m and β, will probably not be uniquely measured from the 2dF QSO survey using this technique. However, all is not lost. There are other ways in which the cosmological parameters can be estimated from the 2dF QSO survey and results can be combined to obtain stronger constraints on cosmology.

The amplitude of QSO clustering can be measured at $z = 1.4$ assuming any cosmology. In order to calculate the value of $\beta(z)$ at $z = 1.4$, we need to know the clustering amplitude of the mass at this redshift. To obtain this, we need to know the amplitude of the clustering of the mass at $z = 0$ and we must assume the cosmological parameters, Ω_m and Ω_Λ. This then allows us to trace the evolution of the mass clustering with redshift in linear perturbation theory.

However, the clustering of the mass at $z = 0$ is not known. Here we use information from galaxy redshift surveys in order to determine the mass correlation function at $z = 0$. The mass correlation function can be determined if the galaxy correlation function and β_g are known, assuming that bias is independent of scale. There are other ways to obtain the mass correlation function at $z = 0$. For example, we could use CDM simulations at $z = 0$ to obtain the mass correlation function amplitude in real space, under assumptions of the shape and normalisation of the CDM models. Alternatively, if the QSO clustering amplitude evolves as a function of redshift, a bias model can be fitted to the QSO clustering and from this, the QSO-mass bias as a function of redshift can be determined, as shown in Croom & Shanks (1996). The latter two approaches are model dependent, whereas the model we describe (once the QSO correlation function has been measured from the 2dF QSO Survey) will be model independent.

7.7.1 Including measurements of β_g

Many measurements of β_g at $z = 0$ have appeared in the literature, for example Tadros & Efstathiou (1996), Ratcliffe et al. (1998c), see also Chapter 3. Early results from the 2dF Galaxy Redshift Survey also provide measurements of β (Peder Norberg, private communication). The value of β_g from optically selected redshift surveys is uncertain but a value of 0.50 ± 0.20 spans all current estimates. For each cosmology (again we only consider flat cosmologies) the value of the galaxy-mass bias can be found from β_g, which in turn gives the value of the mass correlation function if the galaxy correlation function at $z = 0$ is known, as shown below. Rather than determining the value of the two-

point correlation function at one particular point, we use the less noisy volume averaged two-point correlation function, $\bar{\xi}$.

To estimate the redshift space, volume averaged two-point correlation function of galaxies, $\bar{\xi}_g^s$, we assume that the galaxy correlation function can be approximated by a power law of the form $\xi_g^s = (s/6h^{-1}\mathrm{Mpc})^{-1.7}$. $\bar{\xi}_g$ is then found via

$$\bar{\xi}_g^s = \frac{\int_0^{20} 4\pi s^2 (s/6)^{-1.7} ds}{\int_0^{20} 4\pi s^2 ds}. \tag{7.17}$$

By integrating out to $20h^{-1}$Mpc, the non-linear effects on the volume averaged correlation function should be small. The power law approximation is in very good agreement with early results from the 2dF Galaxy Redshift Survey two-point correlation function over the range of scales $2< s <20h^{-1}$Mpc (Peder Norberg, private communication). Once the 2dF Galaxy Redshift Survey is completed, there will be no need to make this approximation.

$\bar{\xi}_q^s$, where the subscript q stands for QSO, is found from the *Hubble Volume* by finding the number of DD, DR and RR pairs that have separations in the range 2-$20h^{-1}$Mpc and forming the two-point correlation using the estimator of Hamilton. This is found for each cosmology and is shown in Table 7.2. The values found for Ω_m=1 and Ω_m=0.3, Ω_Λ=0.7 agree with the values from the January 2000 catalogue as given in Croom et al. (2000).

For each cosmology, the bias between the galaxies and the mass at z=0, $b_{g\rho}(0)$, can be found from

$$b_{g\rho}(0) = \frac{\Omega_\mathrm{m}^{0.6}(0)}{\beta_\mathrm{g}(0)}. \tag{7.18}$$

The real space galaxy correlation function can be determined from the redshift space galaxy correlation function

$$\bar{\xi}_g^r = \frac{\bar{\xi}_g^s}{[1 + \frac{2}{3}\beta_\mathrm{g}(0) + \frac{1}{5}\beta_\mathrm{g}^2(0)]}, \tag{7.19}$$

where the superscripts r and s indicate real and redshift space respectively. The real space mass correlation function at z=0 can now be found as

$$\bar{\xi}_\rho^r(0) = \frac{\bar{\xi}_g^r(0)}{b_{g\rho}^2(0)}. \tag{7.20}$$

The real space mass correlation function evolves according to linear theory such that

$$\bar{\xi}_\rho^r(z) = \frac{\bar{\xi}_\rho^r(0)}{G(z)^2}, \tag{7.21}$$

where $G(z)$ is the growth factor, which depends on cosmology, found from the formula of Carroll, Press & Turner (1992). When Ω_m=1, $G(z) = (1+z)$.

We now relate the correlation function of the mass at z measured in real space to the amplitude of the QSO clustering at z, measured in redshift space as we wish to know $\beta_q(z)$ as a function of $\Omega_{\rm m}(0)$. First we calculate $\Omega_{\rm m}(z)$ using

$$\Omega_{\rm m}(z) = \frac{\Omega_{\rm m}(0)(1+z)^3}{\Omega_{\rm m}(0)(1+z)^3 + \Omega_\Lambda}. \tag{7.22}$$

This is valid for flat cosmologies only. $\beta_q(z)$ is then given by

$$\beta_q(z) = \frac{\Omega_{\rm m}(z)^{0.6}}{b_{q\rho}(z)}, \tag{7.23}$$

but as yet $b_{q\rho}(z)$ is unknown. $b_{q\rho}(z)$ is defined as

$$b^2_{q\rho}(z) = \frac{\bar{\xi}^r_q(z)}{\bar{\xi}^r_\rho(z)} \tag{7.24}$$

substituting the above equations into equation 7.23 gives

$$\beta^2_q(z) = \Omega_{\rm m}(z)^{1.2}\frac{\bar{\xi}^r_\rho(z)}{\bar{\xi}^r_q(z)}. \tag{7.25}$$

$\bar{\xi}^r_q(z)$ can be found via

$$\bar{\xi}^r_q(z) = \frac{\bar{\xi}^s_q(z)}{(1+\frac{2}{3}\beta_q(z)+\frac{1}{5}\beta^2_q(z))}. \tag{7.26}$$

Substituting this into equation 7.25, gives

$$\beta^2_q(z) = \Omega_{\rm m}(z)^{1.2}\frac{\bar{\xi}^r_\rho(z)}{\bar{\xi}^s_q(z)}[1+\frac{2}{3}\beta_q(z)+\frac{1}{5}\beta^2_q(z)], \tag{7.27}$$

a quadratic in $\beta^2_q(z)$ which can easily be solved, allowing $\beta_q(z)$ for any $\Omega_{\rm m}(0)$ to be found. We substitute $z = 1.4$ to find the value of $\beta_q(z)$ at the average redshift of the survey. The values obtained for $\beta^2_q(z)$ are shown in Table 7.2.

The errors on $\beta_q(z)$ are found in the standard way, i.e. by differentiating $\beta_q(z)$ with respect to all the variables that contribute to the error and summing the components of the error in quadrature. In this case there are errors on $\beta_g(0)$, $\bar{\xi}^s_q(z)$ and $\bar{\xi}^s_g(0)$. The error on $\beta_g(0)$ is $\pm$ 0.2, the error on $\bar{\xi}_q(z)$ is the Poisson error found from the total DD counts. The error on $\bar{\xi}_g$ is the least well known as the power law approximation is used. We assume an error of around 20%, which is larger than the error expected from the completed 2dF Galaxy Redshift Survey. The total error on $\beta_q(z)$ is currently dominated by the error on $\beta_g(0)$ in any case. Once the 2dF Galaxy Redshift Survey is complete, $\beta_g(0)$ should be known to better than 10%. The value of $\beta_q(z)$ predicted for each value of $\Omega_{\rm m}$ is given in Table 7.2.

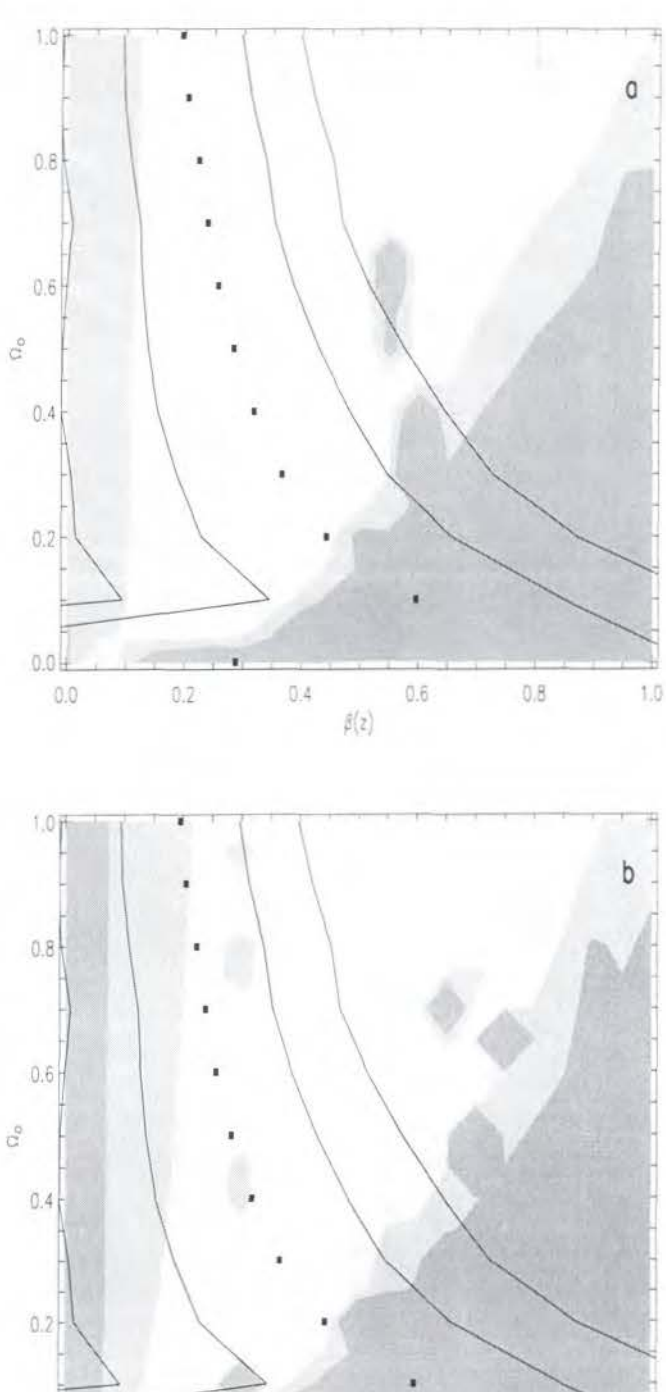

Figure 7.12: *Combining* predicted results from measuring $\xi(\sigma, \pi)$ from the 2dF QSO survey (contour plot as in Figure 7.9) with results from the evolution of clustering. Plot (a) shows the contours with Ω_m=1 assumed and plot (b) shows the contours with Ω_m=0.3, Ω_Λ=0.7. The points show the values of $\beta_q(z)$ predicted from the method assuming β_{gal}. The lines show the 1 and 2σ contours.

$\Omega_m(0)$	$\Omega_m(1.4)$	$\bar{\xi}^r_\rho(0)$	$\bar{\xi}^r_\rho(1.4)$	$\bar{\xi}^s_\rho(1.4)$	$\bar{\xi}^s_q(1.4)$	$\beta_q(1.4)$	Error
1.000	1.000	0.045	0.008	0.016	0.234	0.196	0.101
0.900	0.992	0.051	0.009	0.018	0.249	0.205	0.108
0.800	0.982	0.059	0.011	0.022	0.249	0.223	0.115
0.700	0.970	0.070	0.014	0.027	0.265	0.238	0.118
0.600	0.954	0.084	0.017	0.034	0.283	0.257	0.129
0.500	0.933	0.104	0.023	0.045	0.302	0.283	0.146
0.400	0.902	0.136	0.033	0.064	0.326	0.318	0.164
0.300	0.856	0.192	0.052	0.099	0.360	0.366	0.180
0.200	0.776	0.313	0.099	0.184	0.403	0.444	0.234
0.100	0.606	0.718	0.294	0.512	0.446	0.596	0.331
0.001	0.014	180.4	167.9	180.7	0.456	0.288	0.781

Table 7.2: The values of $\beta_q(1.4)$ predicted from each value of Ω_m assuming β_{gal} at $z = 0$ is 0.50±0.2 and ξ^s_g=0.30±0.06, as described in the text.

7.7.2 Discussion

A comparison between the constraint on Ω_m and $\beta_q(z)$ found here and the constraint found from fitting models to $\xi(\sigma, \pi)$ is shown in Figure 7.12. The contours are shown as in Figure 7.9 and the solid lines show the 1σ and 2σ constraints from the method of Section 7.7. We show the comparison for the two assumed cosmologies. It can be seen that Ω_m=0, Ω_Λ=1 and Ω_m=1, Ω_Λ=0 are ruled out with $\sim 2\sigma$ confidence when the results are combined. The value of Ω_m is now constrained to be $\sim 0.3^{+0.3}_{-0.2}$, by considering where the 1σ contours from the two estimates overlap.

The constraints on $\beta_q(z)$ are stronger than the constraints on cosmology. $\beta_q(z)$ can be constrained from the $\xi(\sigma, \pi)$ contours alone but by combining the errors, $\beta_q(z)$ can be measured to an accuracy of $\sim \pm 0.1$ which puts joint constraints on cosmology and the QSO-mass bias.

The value of Ω_m found here is ~0.3, which is the value used to run the simulation. This shows that the method works under the assumptions made here, i.e. that the cosmology is Ω_m=0.3, Ω_Λ=0.7, that we are working in the linear regime and that the bias is scale independent. The first assumption could be tested if we had different simulations with different cosmological parameters. To accommodate a scale dependent bias, new models would be required and this is the subject of further work.

7.8 Conclusions

We have developed a model for $\xi(\sigma,\pi)$ that takes into account all the different types of distortions that are introduced into the clustering pattern from a redshift survey, such as small scale peculiar velocities, bulk motions and distortions due to the assumed cosmology differing from the underlying cosmology. We have checked that the analytic models are consistent with $\xi(\sigma,\pi)$ measured from the mock *Hubble Volume* catalogues and find that the model fits the simulation well over the range of scales $4 \lesssim \pi, \sigma \lesssim 40h^{-1}$Mpc.

We fit the models to $\xi(\sigma,\pi)$ measured from the *Hubble Volume* and, unfortunately, predict that cosmology will not be strongly constrained from the 2dF QSO Survey. The errors are such that the overall amplitude of the contours dominates the goodness of the fit, rather than the precise shape of the contours.

It is disappointing that the cosmology, $\Omega_{\rm m}$ and Ω_Λ, cannot be directly obtained from the measurement of $\xi(\sigma,\pi)$ from the 2dF QSO survey. However, constraints on $\beta(z)$ are possible from fitting models of $\xi(\sigma,\pi)$ to measurements of $\xi(\sigma,\pi)$ from the mock catalogues. This provides a joint constraint on the cosmology and the QSO-mass bias.

When the results from $\xi(\sigma,\pi)$ are combined with other information, such as the bias between the mass and QSO clustering at $z = 1.4$, constraints on cosmology are possible. If we know the value of $\beta_{\rm gal}$ at $z = 0$, then we can predict the dark matter correlation function at any redshift for any cosmology. Combining results from $\xi(\sigma,\pi)$ and the evolution of clustering provides a constraint that rules out values of $\Omega_{\rm m}$ =0 and $\Omega_{\rm m}$=1 with 2σ confidence, if we start with an $\Omega_{\rm m}$=0.3, Ω_Λ=0.7 simulation.

Simulations with different underlying cosmologies and different models for the bias are required to test just how robust this method of determining cosmology is.

Chapter 8

Distances to Cepheid Open Clusters

8.1 Introduction

Determining the value of Hubble's constant, H_o, has been a challenge to astronomers since the discovery of the expansion of the Universe by Hubble (1929). It is sometimes argued that we are now at the fine tuning stage and many measurements give values for H_o which lie between the hotly argued values of 50km $s^{-1}Mpc^{-1}$ (Sandage) and 100km $s^{-1}Mpc^{-1}$ (de Vaucouleurs), e.g. Tanvir, Ferguson & Shanks (1999) calculated H_o=67±7km $s^{-1}Mpc^{-1}$. However, many of these measurements are based on secondary indicator methods which in turn are dependent on the accuracy of primary indicators of distance, such as the Cepheid Period-Luminosity (P-L) relation. The well-studied LMC P-L relation is usually calibrated via the distance modulus to the LMC and the previously accepted value was around 18.50. However, this has been recently challenged in a paper by Feast & Catchpole (1997) who determined the distance modulus to the LMC as 18.70±0.1. This small difference in the distance modulus causes a 10% decrease in estimates of Hubble's Constant. This discrepancy has further motivated us to check the Galactic zeropoint of the P-L relation. We do this by checking the values of the distance modulus and reddening of a sample of 11 Galactic clusters that contain Cepheids via zero age main sequence fitting (ZAMS).

Previous work on measuring the reddening and distance to young open clusters which contain Cepheids via ZAMS fitting has been done using photoelectric and photographic data in optical wavebands. It is time consuming to observe a large number of stars using photoelectric observations as each star has to be observed individually and photographic data can give relatively inaccurate magnitudes and colours. However, CCD's now make it possible to observe a large number of stars in many different wavebands quickly and accurately. Although CCD's have already been used for open cluster studies e.g. Walker

(1985a), Walker (1985b), Romeo et al. (1989), these have mainly been carried out in *BVRI*. Recently, CCD's with improved *U*-band sensitivity have become available and *U*-band CCD data is included in this study. Infrared imaging detectors are also now available and although some of the detectors used here do not cover as wide an area as optical CCD's, observing the full extent of an open cluster with a mosaic of pointings is a practical proposition.

Until fairly recently, good quality infrared measurements of the Cepheids themselves were not available and the Cepheid P-L relation has been primarily calibrated in the *V*-band. Laney and Stobie (1993,1994) present infrared along with *V*-band magnitudes for a large number of Southern Hemisphere Galactic Cepheids. Using data in the literature to obtain values for the distance modulus and reddening to the clusters they calibrated the Cepheid P-L relation in the *V* and *K*-band. However, any errors in the determination of the distance modulus and the reddening in the previous work would cause an error in the P-L relation as determined by Laney and Stobie. Hence, in this work, we check the previous measurements of the reddening and distance modulus obtained from the open clusters.

The layout of this Chapter is as follows. In Section 8.2 we present the observational data and we test the accuracy of the photometry and calibration of the data. In Section 8.3.1 we describe how the reddenings and distances to the open clusters are obtained and in Section 8.3.2 we discuss each cluster individually. In Section 8.4 we use these values with the magnitudes of the Cepheids to calibrate the Cepheid P-L relation, In Section 8.5 we discuss the implications of the results, particularly for the clusters whose *U-B:B-V* diagrams do not appear to follow the canonical locus. In Section 8.6 we draw conclusions.

8.2 Data

8.2.1 Observations

The observations of the Galactic Open clusters were taken during five observing runs on the JKT, UKIRT, at CTIO, at Calar Alto and on the WHT over a two year period. The spread in declination of the clusters and the multi-wavelength nature of the study meant that many different telescopes were required.

Cluster	Pointing
NGC6649	Star 19 in Madore & van den Bergh (1975)
M25	Star 95 in Sandage (1960)
NGC6664	Star 5 in Arp (1958)
WZ Sgr	WZ Sgr
Lynga 6	TW Nor
NGC6067	Star 136 in Thackeray et al. (1962)
vdBergh 1	CV Mon
TR35	5″ south of TR35
NGC6823	Star j in Guetter (1991)
NGC129	Star 113 in Arp et al. (1959)
NGC7790	Star E in Romeo et al. (1989)

Table 8.1: Approximate pointings for the clusters in the study in all wavebands.

JKT

Optical imaging of eight open clusters was obtained during an observing run from the 16/9/1997 to the 22/9/1997. The observations were carried out using the 1024×1024 Tektronix CCD with pixel scale of 0.33 arcsec pixel^{-1}. Typical seeing was around 1.3″. Short exposures of 5s in V, 10s in B and 20s in U were observed for calibration purposes but the main imaging observations were typically 6x120s in the V-band, 6x180s in the B-band and 6x300s in the U-band. Due to the Southerly declination of some of the objects, they had to be observed at high air mass. However, these observations were normally used to obtain relative photometry and calibration frames were observed at as low an airmass as possible or during a later observing run at CTIO. The pointings are given in Table 8.1. For calibration purposes, standard stars from Landolt (1992) were observed. On the one fully photometric night (21/9/97) six Landolt fields were observed at regular intervals throughout the night, with most of these fields containing several standard stars.

CTIO

The observations were made using the CTIO 0.9m during an observing run from the 24/9/98 to 29/9/98. These observations were carried out using the 2048x2048 Tek#3 CCD with pixel scale 0.384 arcsec pixel^{-1}. The Tek#3 CCD has low readout noise (4 electrons) and good quantum efficiency in the U-band. The average seeing during

the observations was around 1.4″. Approximate pointings are again given in Table 8.1. Standard E-region fields from Graham (1982) and standard stars from Landolt (1992) were observed throughout the night. The only photometric night was the 28/9/1998 and on this night six E-region standard fields and one Landolt standard field containing four standard stars were observed. Three new clusters were observed at CTIO and further observations of clusters observed at the JKT were made in cases where the clusters had been observed in non-photometric conditions only.

UKIRT

The infrared data was mostly observed at UKIRT during the four nights 16-19/6/97 using the IRCAM3 near-IR imaging camera with a 256x256 detector. The pixel scale used was 0.286 arcsec pixel^{-1} giving a field of view of 73″. To cover a sufficient area of each open cluster we therefore had to mosaic images. Generally a mosaic of 9x7 images was observed. Each image was overlapped by half in both the x and y direction so the final image had an approximate area of 6′× 5′. Observations were generally taken using the ND-STARE mode, where the array is reset and read immediately and then read again after the exposure which reduces the readout noise to 35e^{-}. The exposures were 60×2 seconds and the centre of the mosaic is approximately in the same position as the corresponding optical frame. Standards from the UKIRT faint standards list were observed throughout the nights. Around 10 standards were observed on the three photometric nights, some of which were observed early on in the night, half way through and at the end of the night. The seeing throughout the run was typically 0.6″. Observations were also taken in the J and H-band.

Calar Alto

NGC129 lies further north than the declination limit of UKIRT so infrared observations were instead taken at Calar Alto during another observing run. Observations were done using the Rockwell 1k×1k Hawaii detector with pixel scale 0.396 arcsec pixel^{-1}. This gives a 6.′6″ square field of view so there was no need for the mosaicing technique used at UKIRT. The exposure time was 10×1.5s, the seeing was better than 1″ and the exposure was centred on star 113 in Arp et al. (1959). Observations were made in the K_{short}-band. UKIRT faint standards and standards of Hunt et al (1998) were observed for calibration.

Telescope	Date	Cluster	Cepheid	Wavebands	Airmass	Exposure Time(s)	Photometric
JKT 1.0m	17/9/97	NGC6823	SV Vul	*UBV*	1.22	2x150, 2x150, 2x150	No
JKT 1.0m	17/9/97		WZ Sgr	*UB*	1.71	3x150, 2x150	No
JKT 1.0m	18/9/97	M25	U Sgr	*UBV*	1.61	2x150, 1x150, 1x150	No
JKT 1.0m	18/9/97	NGC129	DL Cas	*UBV*	1.36	4x180, 2x180, 1x180	No
JKT 1.0m	20/9/97	NGC6649	V367 Sct	*UBV*	1.39	4x300, 2x180, 2x120	No
JKT 1.0m	21/9/97	NGC6664	EV Sct	*UBV*	1.66	6x300, 8x300, 10x120	Yes
JKT 1.0m	21/9/97	Trumpler 35	RU Sct	*UBV*	1.22	1x180, 1x90,1x90	Yes
JKT 1.0m	21/9/97	NGC7790	CEa, CEb, CF Cas	*UBV*	1.29	6x300, 8x180, 10x120	Yes
CTIO 0.9m	24/9/98		WZ Sgr	*UBV*	1.12	5x300, 4x150,4x90	No
CTIO 0.9m	27/9/98	NGC6067	V340 Nor, QZ Nor	*UBV*	1.36	3x300, 1x90, 1x90	No
CTIO 0.9m	27/9/98	NGC6649	V367 SCT	*UBV*	1.26	5x600, 1x60, 1x60	No
CTIO 0.9m	28/9/98	Lynga 6	TW Nor	*UBV*	1.51	4x300, 1x90, 1x90	Yes
CTIO 0.9m	28/9/98	NGC6067	V340 Nor, QZ Nor	*UBV*	1.36	3x300, 1x90, 1x90	Yes
CTIO 0.9m	28/9/98	vdBergh 1	CV Mon	*UBV*	1.26	1x300, 1x90, 1x90	Yes
CTIO 0.9m	28/9/98	NGC6649	V367 SCT	*UBV*	1.26	1x300, 1x60, 1x30	Yes
CTIO 0.9m	28/9/98	M25	U Sgr	*UBV*	1.27	1x60, 1x5, 1x5	Yes
UKIRT 3.8m	16/6/97	NGC6649	V367 Sct	K	1.40	60x2	Yes
UKIRT 3.8m	17/6/97	M25	U Sgr	K	1.74	60x2	Yes
UKIRT 3.8m	17/6/97	Trumpler 35	RU Sct	K	1.64	60x2	Yes
UKIRT 3.8m	19/6/97	NGC6664	EV Sct	K	1.14	60x2	Yes
UKIRT 3.8m	19/6/97	NGC6823	SV Vul	K	1.03	60x2	Yes
Calar Alto 3.5m	18/8/97	NGC129	DL Cas	$K_{\rm short}$	1.25	10x1.5	Yes
WHT 4.2m	1/9/96	NGC7790	CF, CEa, CEb Cas	$K_{\rm short}$	1.205	50x1	Yes

Table 8.2: Details of the observations. The airmass is the average airmass of the exposure and the exposure time is given in seconds for each of the wavebands in column 5.

WHT

NGC7790 also lies further north than the UKIRT declination limit and so infrared observations were made on the 4.2m WHT during another observing run. The observations were made on the 1/9/1996. The WHIRCAM 256×256 detector which was situated at the Nasmyth focus and behind the MARTINI instrument was used for the observations, without MARTINI tip-tilt in operation. The WHIRCAM detector was the IRCAM detector previously used at UKIRT. The pixel size was 0.25 arcsec pixel^{-1} and the field-of-view was therefore 64″. The observations were centred on star E in Romeo et al. (1989). The $K_{\rm short}$ filter was used and UKIRT faint standards were observed for calibration.

All the observations are summarized in Table 8.2. The date of each observation, the wavebands observed for each cluster and the airmass are given in columns 2, 5 and 6 respectively. Column 7 gives the exposure time of the frames used for imaging. For some of the clusters a calibration frame was observed at CTIO and the exposure time of these clusters are also given in column 7. Column 8 indicates where a cluster was observed on a photometric night and hence where an independent zero point was obtained.

8.2.2 Data Reduction

JKT

Removal of the bias introduced into the data and trimming of the frames to remove the overscan region was done on all the frames using the IRAF task CCDPROC. At least eight U-band sky flats and six B and V-band sky flats were observed on each of the nights so a separate flat field was created for every night using a combination of dust and dawn sky flats. This was created within FLATCOMBINE using a median combining algorithm and a 3σ clipping to remove any cosmic rays. The residual gradient in the flat fields is only around 1%. The task CCDPROC then applies the flat fields to all the images. The same flat fields were used in the reduction of the standard star frames.

Many images of the same cluster were observed. These were all combined together by aligning the images with linear shifts using the task IMSHIFT. Generally these shifts were small (a few pixels either way) as the observation were done one after each other and in some cases no shifts were required. The images were combined using IMCOMBINE and were averaged together using a 5σ clipping.

CTIO

The data was obtained at CTIO using the four amplifier readout mode. A package called QUADPROC within IRAF corrects for the different bias levels in the four quadrants of the CCD. The frames were also trimmed using QUADPROC. In a previous observing run, Croom et al. (1998b) had found that there was a residual gradient of 5% in the dome flat fields so sky flats were used. At least three sky flats were observed on each night in the B and V-band and at least 5 sky flats were observed on each night in the U-band. The resulting flat fields were flat to 1%. The equivalent version of FLATCOMBINE in the QUAD package was used to median combine the flat fields using a 3σ clipping. The E-region standards and observations of Landolt standards were again reduced in the same manner.

Multiple images of the same cluster were again combined using IMCOMBINE with the same settings as for the JKT data and where any offset shifts appeared between the data frames they were again corrected for using IMSHIFT.

UKIRT

The UKIRT data was reduced using a program called STRED within the package IRCAMDR. This is a fairly automated routine which reads in the data frames, subtracts of the dark count and creates a flat field frame by median filtering the image frames. Then the program flat fields the dark subtracted object images, corrects for any bad pixels and finally creates a mosaic. All the data frames were median combined to create a flat field for each night. There was no evidence of a large scale gradient greater than about 1% in the flat fields.

To create the final image, all the individual frames have to be mosaiced together. STRED reads in the offset from the data header, however, these offsets were not accurate enough. By creating a separate offsets file the mosaicing could be done more accurately. To create the offsets file, one of the corner frames was fixed and the offset required for the neighbour frame were found by eye. This was built up over the whole frame, however once one offset had been determined, all the other offsets in the x and y direction from frame to frame were the same. The offsets for the standard stars were more accurate and could be used to create the mosaic.

Calar Alto

Basic IRAF routines such as IMCOMBINE and IMARITH were used to reduce the Calar Alto data. A flat field was created by median combining all the data frames. A more detailed description of the data reduction can be found in McCracken (1999). We found that the best results were obtained by subtracting a sky frame from each image, as follows: A sky frame for each individual image was created using IMCOMBINE with a 5σ clipping to median filter four data frames that were local in time to the image frame to create a sky frame for each data frame. The sky frame was then subtracted off the image before IMCOMBINE was used again to combine all the data frames, forming one final image frame.

WHT

The data from the WHT was observed and reduced for us as part of another project. Dome flats were used to flat field the data as no sky flats were available and not enough data frames were taken to create a flat frame by median combining the data frames. However when this is divided into the science frame using IMARITH, the resulting data frame appears fairly flat. Sky subtraction was also required. A sky frame was created by combining dedicated sky frames observed locally in time to the science frame and this was then subtracted from the science frame using IMARITH.

8.2.3 Image Alignment

In order to be able to produce colour-magnitude diagrams, the magnitude of each star in all the different wavebands is required. To do this, all the frames need to be aligned. Aligning the optical frames was easy as there were only linear shifts between each waveband. Rather than altering the data, the alignment was just done by applying small corrections to the x and y positions of the stars. Aligning the optical data with the K-band data was more complicated though as there were shifts, shears and rotations between the frames. We used the IRAF routine GEOMAP to calculate the best spatial transformation function between any two images thus allowing the optical and K-band data to be aligned. The mapping was only used to transform coordinates and we did not perform photometry on the resampled images.

8.2.4 Photometric Calibration

JKT

Observations of the standard stars of Landolt (1992) were taken at regular intervals on each of the six nights at the JKT. The CCD frames of the standard stars were reduced in the same manner as the data frames with the same flat fields etc as discussed in Section 8.2.2. The aperture size used to measure the magnitude of the standard stars was 5″ as this was the size required when measuring the magnitudes of the stars on the data frames to avoid any problems with crowding.

Only the sixth night (21/9/97) was fully photometric and the zero point, airmass coefficient and colour equation are shown in Figure 8.1 and given below with the rms scatter.

$$U_{\rm jkt} = U_{\rm ldt} + 4.03 + 0.42sec(z) - 0.062(U_{\rm ldt} - B_{\rm ldt}) \pm 0.039$$

$$B_{\rm jkt} = B_{\rm ldt} + 2.17 + 0.23sec(z) - 0.034(B_{\rm ldt} - V_{\rm ldt}) \pm 0.037$$

$$V_{\rm jkt} = V_{\rm ldt} + 2.14 + 0.17sec(z) \pm 0.029$$

where the subscript ldt stands for the Landolt standard star magnitude and the subscript jkt stands for the instrumental magnitude, z represents the zenith distance. The errors on the airmass are ±0.0025, 0.0012 and 0.0008 in the U, B and V-bands respectively and the errors on the colour equation are ±0.0007 and 0.0005 in the U and B-bands respectively. As mentioned above, a 5″ aperture was used to measure the magnitudes of the standard stars. The zero point was also checked with a 15″ aperture and the colour equation and airmass terms were found to be very similar. The only difference between the two calibrations was the value of the zero point correction as more light is detected with the larger 15″ aperture.

There are no Landolt magnitudes available for the data frames so the colour term has to be translated into instrumental magnitudes. The colour term is negligible in the V-band calibration so the instrumental V-band magnitudes come directly from the above equations. The $B_{\rm ldt}$-$V_{\rm ldt}$ and $U_{\rm ldt}$-$B_{\rm ldt}$ colours are given by

$$(B_{\rm ldt} - V_{\rm ldt}) = 1.035((B_{\rm jkt} - V_{\rm jkt}) - 0.03 - 0.06sec(z))$$

$$(U_{\rm ldt} - B_{\rm ldt}) = 1.067((U_{\rm jkt} - B_{\rm jkt}) - 1.86 - 0.19sec(z))$$

we assume that the contribution from the colour term in the calibration of the B-band is negligible when determining the $U_{\rm ldt}$-$B_{\rm ldt}$ colour as compared to the offset and airmass

terms. For a star with B_{jkt}-$V_{jkt} \sim 2$ then the difference in the U-band magnitude with and without the B_{ldt}-V_{ldt} term is 0.004 magnitudes. These colour terms are used to correct the instrumental magnitudes in order to make the colour-colour and colour-magnitude diagrams in Figures 8.14, 8.15, 8.16, 8.17 and 8.18.

The clusters NGC7790, NGC6664 and Trumpler 35 were observed on the one photometric night. Short exposure observations of NGC6649 and M25 were made at CTIO in order to obtain an independent zero point for these frames. Around 20 bright (brighter than $V\sim15$), fairly uncrowded, unsaturated stars were taken as standard stars to identify the relation between the zero point from the JKT data and from CTIO data to an accuracy of a few hundredths of a magnitude. The agreement between the zero point found via this method and previous, photoelectric calibrations is good (see Table 8.3) with only small offsets in each case.

The remaining three clusters were observed in non-photometric conditions only so previous work had to be relied upon for the calibration. For NGC6823 the photoelectric observations from Table 1 of Guetter (1991) were used. Some of these stars were saturated on the CCD frame and the area of overlap between the two images was not identical but thirteen stars were suitable for calibration purposes. Guetter (1991) compares his photoelectric data with that of previous work and finds good agreement. Two sources of photoelectric data are available for the cluster NGC129, Arp et al. (1959) and Turner, Forbes & Pedreros (1992). There are 9 stars in common with the Arp photometry and 13 stars in common with the Turner photometry. For these samples of stars, we find that the U-band zero point obtained from Arp is 0.04±0.03 mags brighter than that of Turner. In the B-band the difference is 0.02±0.01 mags in the sense that Arp is brighter than Turner and there is 0.01±0.01 mag difference in the same sense in the V-band. These differences are mainly caused by the stars in the sample with V fainter than 14 mag. We therefore use the average of the brightest two stars from Arp and Turner to calibrate NGC129. The zero point from this method agrees very well with the zero point obtained using Turner's photometry, which is shown in Figure 8.21. Finally the photoelectric work of Turner (1984) was used to calibrate the cluster containing the Cepheid WZ Sgr.

CTIO

E-region standards from Graham (1982) and Landolt standards were observed for the photometric calibration of the optical CTIO data. Menzies et al. (1991) compared the zero points and colour differences found from using the two different standard star studies and

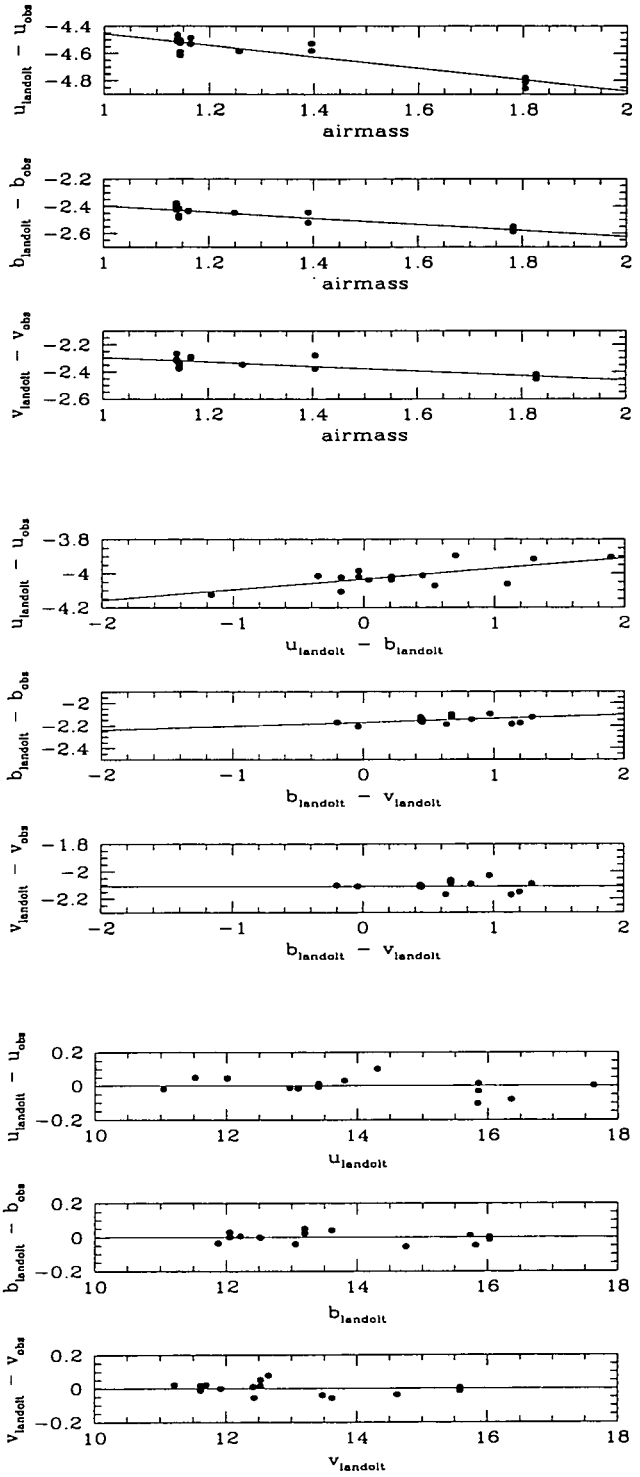

Figure 8.1: The airmass coefficient, colour equation and zero point, with the colour equation and airmass correction applied, for the U, B and V wavebands for the data from JKT.

found the offsets between the two to be small, 0.004±0.0095 offset in the sense E regions - Landolt and similar sized offsets in the *U-B* and *B-V* colours. Any offsets are within the quoted error. As before, these standard frames were reduced in the same manner as the data frames and a 5″ aperture was used to determine the magnitude. Again, only one night was photometric and this was the last night (28/9/98). The zero points, airmass coefficients and colour equations for each waveband, *U*, *B* and *V*, are shown in Figure 8.2 and given below again with the rms scatter.

$$U_{\rm ctio} = U_{\rm std} + 4.74 + 0.47sec(z) - 0.036(U_{\rm std} - B_{\rm std}) \pm 0.037$$

$$B_{\rm ctio} = B_{\rm std} + 3.34 + 0.21sec(z) + 0.099(B_{\rm std} - V_{\rm std}) \pm 0.021$$

$$V_{\rm ctio} = V_{\rm std} + 3.09 + 0.12sec(z) - 0.018(B_{\rm std} - V_{\rm std}) \pm 0.009$$

where the subscript std stands for the E-field or Landolt standard magnitude and ctio stands for the instrumental magnitude. The errors on the airmass are ±0.003, 0.0011 and 0.0009 in the *U*, *B* and *V* and the errors on the colour equation are ±0.0006, 0.0007 and 0.001 in the *U*, *B* and *V*-bands respectively. There is generally good agreement between the values for the airmass coefficients and colour equations found in this work and in Croom et al. (1998b). Again, for the data frames the colour terms have to be found in terms of CCD magnitudes rather than standard magnitudes. The colour term in the *B*-band is in this case non-negligible so *V*-band magnitudes have to be used in the *U*-band calibration. The standard colours are given by

$$(B_{\rm std} - V_{\rm std}) = 0.78((B_{\rm ctio} - V_{\rm ctio}) - 0.25 - 0.09sec(z))$$

$$\begin{aligned}(U_{\rm std} - B_{\rm std}) = 1.037((U_{\rm ctio} - B_{\rm ctio}) - 1.40 - 0.26sec(z) \\ +0.099(0.78((B_{\rm ctio} - V_{\rm ctio}) - 0.25 - 0.09sec(z))))\end{aligned}$$

again these colours are used to calculate the instrumental magnitudes which are used in the colour-colour and colour-magnitude diagrams in Figures 8.14, 8.15, 8.16 8.17 and 8.18.

This night provided independent zero points for the clusters NGC6067, Lynga 6 and vdBergh1, which were not observed at JKT, and also provided a zero point for NGC6649 and M25, which were only observed in non-photometric conditions at the JKT.

UKIRT

The standard stars observed at UKIRT were taken from the faint standards list available from the UKIRT Web page. The standard stars were observed as a mosaic of five frames,

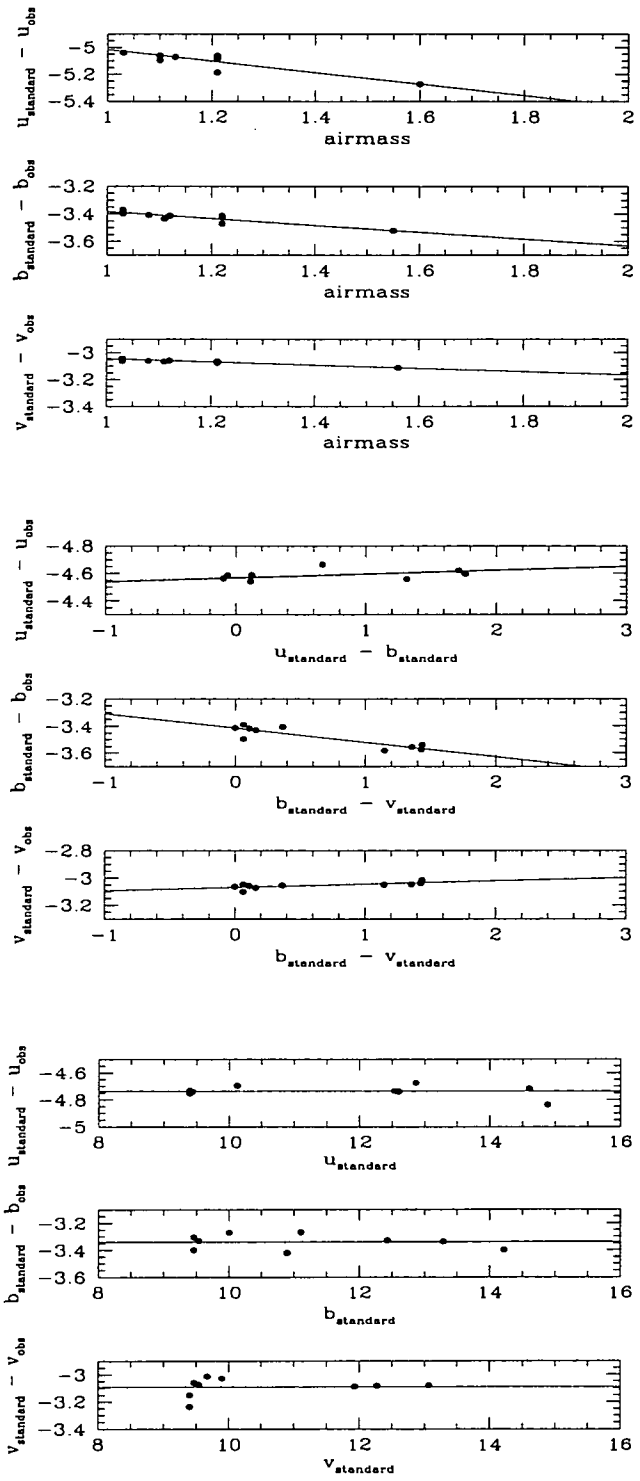

Figure 8.2: The airmass coefficient, colour equation and zero point, with the colour equation and airmass correction applied, for the U, B and V wavebands for the data from CTIO.

a central frame with an overlapping frame in each direction. A 5″ aperture was used to determine the magnitude of the standard. Three of the nights were photometric, night 1, 2 and 4 (16, 17,19/6/97) so all the clusters were observed in photometric conditions.

The zero points for each of the nights are shown in Figure 8.3 and the airmass coefficient and colour equation are shown in Figure 8.4 and given below for each night.
16/06/1997;

$$K_{\rm U} = K_{\rm std} + 7.17 - 0.082sec(z) + 0.005(J_{\rm std} - K_{\rm std}) \pm 0.026$$

17/06/1997;

$$K_{\rm U} = K_{\rm std} + 7.26 - 0.082sec(z) + 0.005(J_{\rm std} - K_{\rm std}) \pm 0.035$$

18/06/1997;

$$K_{\rm U} = K_{\rm std} + 7.27 - 0.082sec(z) + 0.005(J_{\rm std} - K_{\rm std}) \pm 0.038$$

where again the subscript std refers to the standard stars. The difference between each of the calibrations is just a small change in the zero point. The error on the airmass coefficient is ±0.002 and the error on the colour term is ±0.0003. The airmass coefficient agrees well with the values given in Krisciunas et al. (1987) and those found on the UKIRT web page. As the colour term is very small, it was assumed negligible in the *V-K:V* CMD's.

Calar Alto and WHT

UKIRT faint standards and Hunt et al. (1998) standards were observed in order to calibrate the Calar Alto and WHT data. The calibration was provided for us by Nigel Metcalfe as part of another project (see McCracken, Metcalfe and Shanks, 2000, in preparation for more details).

8.2.5 Photometry

Automated aperture photometry was done using PHOT within IRAF's DAOPHOT package. A small, 5″, aperture was used to minimise any crowding problems. PHOT was also used to obtain the magnitudes of the standard stars, the magnitude of each standard star was measured individually using a 5″ aperture.

First we establish the depth of the photometry. We define the limiting depth of the observations to be where the Poisson error in the electron counts is 5%, i.e. when

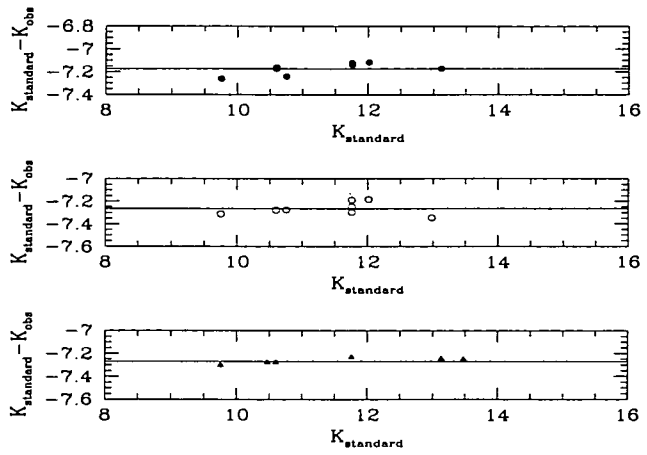

Figure 8.3: The K-band zero points for the three photometric nights at UKIRT. The top panel is for night 1, (16/8/97, filled circles), centre for night 2 (17/8/97, open circles) and the lower panel for night 4 (19/8/97, filled triangles).

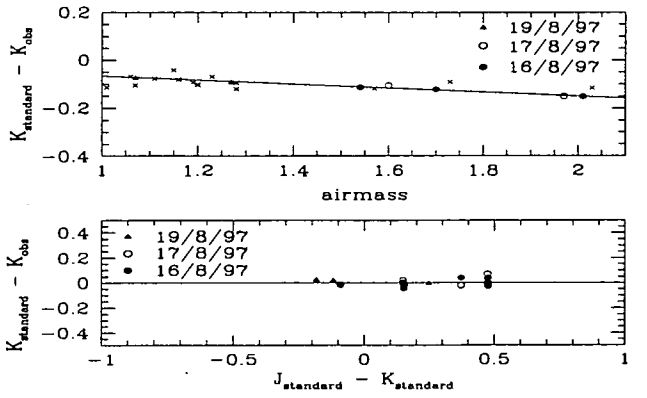

Figure 8.4: The airmass correction and colour equation for the K-band data taken at UKIRT. The symbols are the same as in Figure 8.3 but the crosses in the top panel show stars observed on any one of the nights observed at one airmass only.

Cluster	U_{obs} - $U_{previous}$	B_{obs} - $B_{previous}$	V_{obs} - $V_{previous}$	N(stars)	Previous Work
NGC6649	0.02±0.025	0.01±0.01	0.00±0.012	20	Madore & van den Bergh (1975)
M25	0.01±0.03	-0.02±0.02	-0.005±0.02	21	Sandage (1960)
NGC6664	0.02±0.012	0.01±0.01	-0.015±0.015	15	Arp (1958)
Lynga 6	-0.015±0.02	-0.01±0.02	-0.01±0.015	17	van den Bergh & Harris (1976)
NGC6067	0.07±0.08	0.04±0.03	-0.06±0.04	6*	Thackeray et al. (1962)
vdBergh 1	-0.01±0.015	0.00±0.02	-0.005±0.008	24	Arp (1960)
Trumpler 35	-0.02±0.02	-0.01±0.015	-0.005±0.015	15	Hoag et al. (1961)
NGC7790	-0.009±0.01	0.015±0.008	-0.015±0.007	22	Sandage (1958)
WZ Sgr	0.00±0.01	0.00±0.011	0.00±0.007	12	Turner (1984)
NGC6823	0.00±0.01	0.00±0.005	0.00±0.004	13	Guetter (1991)
NGC129	0.00±0.03	-0.005±0.01	-0.008±0.01	13	Turner, Forbes & Pedreros (1992)

Table 8.3: Comparison of the photometry of this study with previous photoelectric data. Note there are 6 stars from Thackeray et al. (1962) in common with the B and V-band data but only 2 in common with the U-band data. The last three clusters were observed on non-photometric nights only so the previous work was relied upon for calibration. There is a small offset between the zero point of NGC129 and Turner, Forbes & Pedreros (1992) as only the brightest stars were used for calibration.

$\sqrt{N_{obj} + N_{sky}}/N_{obj} = 0.05$. The depth of the JKT data is 19.9 mags for a 1200s U-band exposure, 19.7 mags for a 360s B-band exposure and 20.4 mags for a 240s V-band exposure. The depths of the CTIO data are similar. The depth of a typical K-band exposure of 120s observed was around 19.2 magnitudes.

The accuracy of the zero points were tested by comparing our photometry to photoelectric observations in the literature. Table 8.3 shows the residual when our magnitudes for typically 10 stars from the calibrated data frames (see Section 8.2.4) are compared to the magnitudes found from previous photoelectric studies. In all cases except for NGC6067, the residual is small (less than 0.03 mags). Unfortunately for the cluster NGC6067 there is little photoelectric data available and only 6 stars can be compared in the B and V-bands due to saturation and different parts of the cluster being observed. In the U-band there are only 2 stars in common. Given that the offsets between the work here and previous work are small for all the other clusters, we assume that the zero point obtained for NGC6067 is accurate.

Finally, we test how the crowding of the field effects the photometry. Twenty 'simulated stars' of each magnitude shown in Figure 8.5 were placed in a 300s U-band image of the cluster M25 then PHOT was used to determine how well the magnitudes could be recovered. Shown in Figure 8.5 are the results. The x axis shows the true magnitude of the stars and the y axis shows the deviation of the mean magnitude from the true magnitude. This plot shows the total error (Poisson errors, read noise and also errors due to crowding) at each magnitude (full errorbar) and the Poisson error, found as described above, as the smaller, wider, errorbar. The total error for this exposure of 300s is less than 0.03 magnitudes down to U=17 mags. For a U-band exposure of 1800s, assuming that the crowding errors remain the same and correcting the error shown in Figure 8.5 for the reduced Poisson error, the total error is estimated to be $\pm$0.17 at U=20, reducing to $\pm$0.08 at 18th mag and $\pm$0.03 at 17th mag. Figure 8.5 also shows that although the errors increase for fainter magnitudes there is no systematic trend.

8.3 Reddening and Distance

8.3.1 Method

In order to work out the Cepheid Period-Luminosity (P-L) relation, the distance to the cluster needs to be known. However, there is generally significant dust absorption along the line of sight to the cluster which must be corrected for. The method for determining

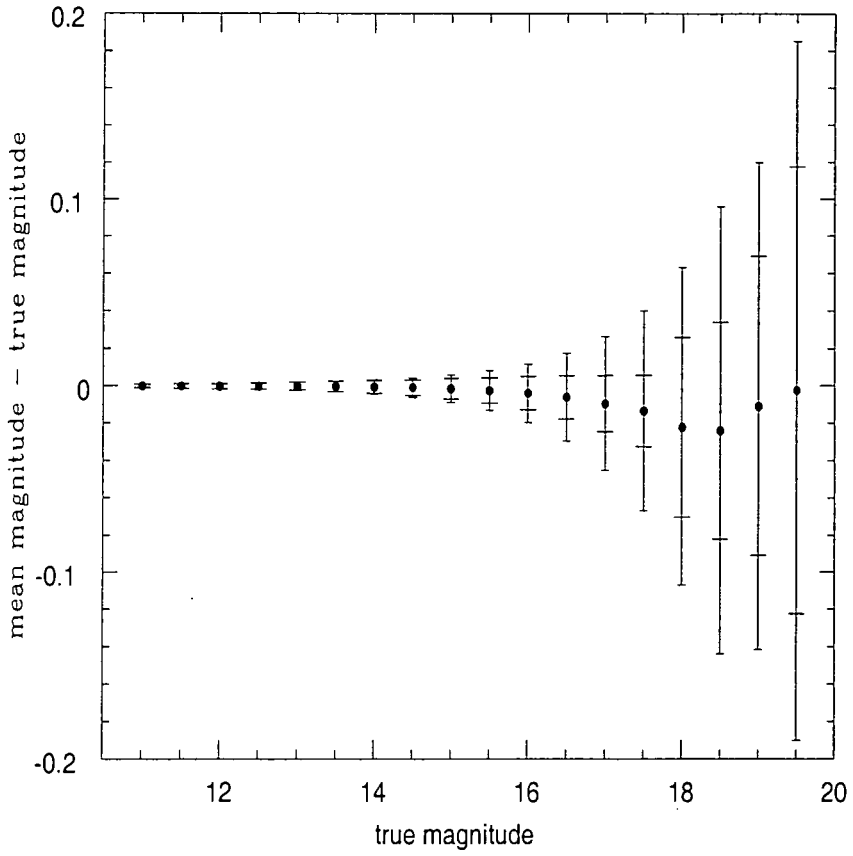

Figure 8.5: An indication of the accuracy of the U-band photometry based on results for simulated stars (see text for details). The large errorbar indicates the total error (Poisson errors, crowding errors and also read noise), the smaller, wider errorbar indicates the Poisson error found as described in Section 8.2.5, paragraph 2. The total error at 17th magnitude is only ± 0.036 magnitudes and down to 18th magnitude the error is less than ± 0.1 magnitudes.

both of these parameters is done via ZAMS fitting to colour-magnitude and colour-colour diagrams. The ZAMS used in this study is a combination of Allen (1973) for the optical data, the intrinsic colours of near IR-band stars from the UKIRT Web page (Tokunaga 1998) for the K-band data and nearby local stars taken from the Strasbourg Catalogue to give an indication of the acceptable spread in the ZAMS. The ZAMS of Turner (1979b) and Mermilliod (1981) have also been tested and would give the same results as the ZAMS of Allen (1973). The reddening law assumed in this work is from Sharpless (1963).

$$\frac{\mathrm{E}(U-B)}{\mathrm{E}(B-V)} = 0.72 + 0.05\mathrm{E}(B-V) \tag{8.1}$$

The reddening is obtained using the *U-B:B-V* diagram which is independent of the distance to the cluster. The *U-B:B-V* diagram has always been the favoured method for estimating the reddening; however, previously, the accuracy of the reddening determination was limited by the depth of the U-band photoelectric photometry. The improved U sensitivity of the current generation of CCD's should allow a potential improvement in the accuracy of the reddening estimated from *U-B:B-V* diagrams and this is the route we have adopted here.

We choose to fit the ZAMS to the ridge-line of the O and B stars rather than the least reddened envelope, because it helps take account of differential reddening in some of the clusters. There is evidence for differential reddening in NGC6823 and TR35 (see Figure 8.15) as the main sequence in the colour-colour diagram is substantially broadened. By fitting to the centre of the data we measure the average reddening for the cluster which we can then apply to colour-magnitude diagrams which are uncorrected for differential absorption to obtain distances.

To determine the error in the reddening we measure the standard deviation of the O and B type stars from the ZAMS via least squares fitting. The errors quoted on the values for the reddening (see Table 8.4) are typically less than 0.1 mags and include any error in the calibration. In quite a few cases, the ZAMS does not fit the colour-colour data well over the whole range of B-V colours. This is particularly problematic in some cases and these cases are discussed below in Section 8.3.2. There is also the problem that the Cepheid could have a different reddening to the cluster, caused by differential reddening across the cluster or by the location of the Cepheid away from the cluster. This is discussed further in Section 8.4.

We use both the *B-V:V* and *V-K:V* colour-magnitude diagrams to determine the cluster distance. *V-K:V* diagrams have the advantage that the slope of the ZAMS is

flatter than for *B-V:V*, possibly allowing more accurate distance estimates but *V-K:V* diagrams are available for only 7 of the clusters and the scatter in these diagrams is greater than in the *B-V:V* CMD's, particularly in the case of NGC6664 as there were difficulties aligning the *V* and *K* frames due to a shift in the telescope position half way through creating the *K*-band mosaic for this cluster. There are also few points on the *V-K:V* diagram for NGC7790 due to the small size of the IRCAM detector. This diagram was also made for us before the *V*-band observations for this cluster were available as a test of the feasibility of this project. The *V*-band data therefore comes from Romeo et al. (1989), however there is a good match between our *V*-band data and that of Romeo et al. (see Figure 8.23). The distance modulus which best fits the *B-V:V* CMD around the position of the A0V stars, using the method of least squares, is taken to be the distance modulus, μ_0, of the cluster. This value is found assuming $R = 3.1$ for each cluster. As the value of R varies from paper to paper, we quote the unreddened distance modulus, μ, in Table 8.4 and compare these value with previous values below in Section 8.3.2. The errors on the distance modulus were found by measuring the standard deviation away from the ZAMS of A0V type stars over the range $-0.1 \lesssim B\text{-}V \lesssim 0.1$ in the dereddened ZAMS which covers a range of approximately 4 mags in *V*. The distance modulus found from the *B-V:V* CMD is then checked against the *V-K:V* CMD for the 7 clusters with such a CMD for consistency. In all cases the distance modulus found from the *B-V:V* diagram was consistent with the *V-K:V* CMD within the errors.

No attempt has been made to remove foreground and background stars. Only stars which lie clearly off the main sequence (off in *B-V* by more than 1 mag for example) were removed. There is no clear recipe for how to remove the contaminating stars from the colour-magnitude and colour-colour diagrams, unless the stars are individually dereddened. Then any stars that require a value of the reddening that is much different from the rest of the cluster are considered to be contaminating stars and are removed from the colour-colour and colour-magnitude diagrams. The data shown in Figures 8.14, 8.15, 8.16, 8.17 and 8.18 therefore may contain non-cluster members. The observations in this study only cover a field-of-view of $\sim 6'$ and are pointed at the cluster centre so contamination may not be as much of an issue as for the wider field photographic plates. As an example, we discuss this further for the cluster NGC7790 in Section 8.3.2.

All of the *U-B:B-V* diagrams, the *B-V:V* and the *V-K:V* diagrams are given in Figures 8.14, 8.15, 8.16, 8.17 and 8.18 respectively and our results for the reddening and distances obtained are given in Table 8.4 together with previous results as summarised by Laney

Cluster	$E(B\text{-}V)_{clus}$	$E(B\text{-}V)^{LS}_{clus}$	μ	μ^{LS}
NGC6649	1.37±0.06	1.35	15.52±0.11	15.42
M25	0.49±0.04	0.48	10.57±0.16	10.49
NGC6664	0.66±0.06	0.64	12.80±0.14	12.35
WZ Sgr	0.56±0.08	0.57	12.88±0.14	12.95
Lynga 6	1.36±0.08	1.34	15.50±0.16	15.54
NGC6067	0.37±0.03	0.35	12.31±0.12	12.19
vdBergh1	0.90±0.09	0.77	14.09±0.19	13.60
Trumpler 35	1.19±0.07	0.92	15.00±0.18	14.37
NGC6823	0.85±0.07	0.75	13.52±0.20	14.09
NGC129	0.57±0.05	0.53*	12.67±0.14	12.85*
NGC7790	0.59±0.05	0.64*	14.37±0.13	14.34*

Table 8.4: The values in the above Table come from the work here and from Laney & Stobie (1993). The values marked * come from Feast & Walker (1987) as the clusters NGC129 and NGC7790 are not included in the studies of Laney and Stobie.

& Stobie (1994). All of the clusters are individually discussed below in Section 8.3.2

8.3.2 Discussion of Individual Clusters

NGC6649 has been studied previously by, for example, Madore & van den Bergh (1975) and Walker & Laney (1986). The agreement between the photometry of this study and the photoelectric data of Madore & van den Bergh (1975) is good (see Table 8.3), with only small offsets in the U and B-band. Madore & van den Bergh (1975) find $E(B\text{-}V)$ = 1.37 (no quoted error) for the reddening towards the cluster. The distance modulus μ from Madore & van den Bergh (1975) is 15.4±0.2.

Turner (1981) uses the photometry of Madore & van den Bergh (1975) and that of Talbert (1975) to study NGC6649 and finds the cluster suffers from differential reddening. However for stars close to the cluster centre a value of $E(B\text{-}V)$=1.38 is appropriate. The dereddened distance modulus, μ_o, is found to be 11.06±0.03 when individual stars are dereddened ($R = 3$ was assumed which implies μ=15.2)

Walker & Laney (1986) used U, B and V-band CCD data to study NGC6649. Agreement between the photoelectric data of Madore & van den Bergh (1975) and Walker & Laney (1986) was found to be better than 0.03 mags in the V-band. Walker & Laney (1986) do not measure the reddening of the cluster due to the claims of differential redden-

ing by Turner. Walker & Laney (1986) deredden each star individually to find a distance modulus of 11.00±0.15 (they assumed $R = 3.27$ giving μ=15.4).

Rather than correct for differential reddening, we fit the ZAMS line to the centre of the *U-B:B-V* and *B-V:V* diagrams to try and measure the average values of the reddening. We find E(*B-V*)=1.37±0.06 and μ=15.52±0.11. Both values are consistent with previous work. The distance modulus used in Laney & Stobie (1994) is μ=15.42 and the reddening, E(*B-V*), is 1.35, consistent with the values found here.

Barrell (1980) found the radial velocity of the Cepheid V367 Sct to be -20±6 km s^{-1} and the radial velocity of the cluster NGC6649 to be -14±5 km s^{-1}. This is taken to be evidence for the cluster membership of the Cepheid.

M25 The photometry used in this study and that of Sandage (1960) is compared in Figure 8.19 and 8.20. The agreement in all the wavebands is good, less that 0.03 mags different from that of Sandage.

M25 has a *U-B:B-V* diagram where the ZAMS fit is good over a wide range of *B-V* colours. The value for E(*B-V*) of 0.49±0.04 is in excellent agreement with the work of Sandage (1960) who obtained 0.49±0.05. Johnson (1960) stated that E(*B-V*) lay in the range 0.4 to 0.56 and van den Bergh (1978) obtained E(*B-V*)=0.51±0.01.

The distance modulus obtained for this cluster is μ=10.57±0.16. This is slightly higher than the value of Sandage, who quote μ_o=8.78±0.15 (which corresponds to μ=10.25 as $R = 3$ was assumed by Sandage). The difference in the distance modulus is probably due to where the fit was made, we fit to the A type stars to obtain 10.57±0.43, Sandage's seemed to fit to brighter stars which may lie slightly off the main sequence. Wampler et al (1960) obtained 9.08±0.2, which translates to μ=10.58 as $R = 3$ was assumed and van den Bergh (1978) found μ= 10.7±0.3 which is slightly larger than our value. There is reasonable agreement between our values and those used by Laney & Stobie (1994), as given in Table 8.4.

Feast (1957) studied M25 by measuring radial velocities and spectra for stars within M25. He obtained radial velocities of around 4 km s^{-1} for the Cepheid U Sgr and for 35 stars in the cluster M25 indicating that U Sgr is a member of M25.,

NGC6664 has been studied previously by Arp (1958). The agreement between his photometry and ours is good, with only small offsets between the two data sets as given in Table 8.3.

Figure 8.14 shows that the cluster NGC6664 has an anomalous *U-B:B-V* diagram. This is a clear case where we fit the O and B type stars rather than trying to fit the F and

G type stars as the F and G type stars maybe more affected by metallicity. The value for E(B-V) is then 0.66±0.06 which is slightly larger than 0.6 obtained by Arp (1958). No error is quoted by Arp. In the previous work by Arp, the observations were not deep enough to see if the anomalous shape of the U-B:B-V diagram would have been detected or not. Unfortunately the only previous source of photometry for more than a handful of stars is that of Arp so no other zero point comparisons can be made.

The distance modulus obtained is 12.87±0.14 which is larger than 12.6±0.2, inferred from Arp (1958), who obtained μ_o = 10.8±0.2 assuming R = 3. This is mostly due to the increased value for the reddening estimated here.

There is a discrepancy between the value of the distance modulus used by Laney & Stobie (1994) who quote a reddening of 0.64 but a distance modulus of 12.35. Becker & Fenkert (1971) quote an even lower value of μ=12.25 (with no error) whereas Feast & Walker (1987) use μ=12.7.

The radial velocity work of Kraft (1958) shows EV Sct is a member of NGC6664.

WZ Sgr The first problem with WZ Sgr is that its membership of an open cluster is questionable. Turner (1984) discusses the membership of WZ Sgr to the cluster C1814-190 in some detail and concludes that the strongest evidence for membership of WZ Sgr to an open cluster comes from the fact that often the Cepheid in a cluster is around 4 magnitudes more luminous than the B-type stars on the main sequence. This essentially means that the age of the Cepheid is consistent with the age of the cluster. Feast & Walker (1987) assume WZ Sgr to be a cluster member.

The cluster C1814-190 is also only comparatively sparsely populated with only around 35 members brighter than B∼16 Turner (1984). There is also patchiness in the dust obscuration which would cause differential reddening (Turner 1984, 1993a). These two factors could possibly explain the slightly odd shape of the U-B:B-V colour-colour diagram particularly in the range $0.5 \lesssim B\text{-}V \lesssim 1$. Contamination from foreground and background stars could also be the source of the unusual U-B:B-V diagram. By measuring the reddening of the O and B type stars we obtain a value for E(B-V)=0.56±0.08 which is in agreement with Turner (1984) and consistent with E(B-V)=0.57 used by Laney & Stobie (1994).

The B-V:V diagram is surprisingly tight, giving a distance modulus of μ=12.88±0.14 which is again in agreement with Turner (1984) who obtained μ_o=11.16±0.1 with R=3.1 implying μ=12.90. The distance modulus used by Laney & Stobie (1994) is similar at 12.95.

Lynga 6 The photometry of Lynga 6 has been checked against that of van den Bergh & Harris (1976) (shown in Figure 8.19 and 8.20) and also against that of Moffat & Vogt (1975). Offsets between the different sets of photometry are less than 0.03 mags in each waveband.

Like NGC6649, Lynga 6 is also very heavily reddened which makes it very difficult to observe stars over a large range of B-V and U-B colours. E(B-V) is estimated to be 1.36±0.08 which is consistent with previous measurements by van den Bergh & Harris (1976) who obtained 1.34±0.01, by Madore (1975) who obtained 1.37±0.03 and the value of 1.34 used by Laney & Stobie (1994).

The distance modulus obtained here is around 15.32±0.16 which is consistent with 15.37 inferred from Walker (1985a) (μ_o=11.15±0.3 found with R=3.15) but is slightly lower than the value used by Laney & Stobie (1994) who quote a distance modulus of 15.54. Madore (1975) obtained a value of μ=16.2±0.5 but only 6 stars were used to obtain this value. The difference between the value of Madore and the value for the distance found here is that Madore tended to fit the edge of the ZAMS.

The Cepheid TW Nor lies close to the centre of the cluster Lynga 6 and has a very similar value of the reddening. This is taken as evidence of the membership of the Cepheid to the cluster (Walker 1985a).

NGC6067 There is very little photoelectric data for this cluster. The B and V-bands have been compared to the data of Thackeray et al. (1962) but there are only 7 stars in common and in the U-band there are only two stars in common. The comparison shows that the zero point used here is at least consistent with previous work. As the photometry obtained at CTIO for other clusters such as Lynga 6 and vdBergh1 agrees well with previous results, we have to assume that the photometry for NGC6067 is also good.

NGC6067 has the lowest value for the reddening of the clusters in this study with E(B-V) estimated as 0.37±0.03. The U-B:B-V diagram presented here has a main sequence which agrees fairly well with the ZAMS, although there is a spread around B-V=0.8. The value of 0.37 is slightly higher than previous values, Coulson & Caldwell (1984) obtained 0.35±0.1 and Thackeray et al. (1962) obtained 0.33 with no quoted error.

The distance modulus is estimated to be 12.31±0.12 which is slightly larger than the value of 12.15±0.1 inferred from Walker (μ_o=11.05±0.1 and $R \sim 3.15$). However Thackeray et al. (1962) found the distance modulus of the cluster to be around 12.5. This estimate is higher than our estimate, despite a smaller measured reddening, as the ZAMS

fit was made to the edge of the CMD.

Laney & Stobie (1994) use 0.35 for the reddening and 12.19 for the distance modulus to the cluster.

The membership of the Cepheids to this cluster is discussed in detail by Eggen (1983). Eggen noted that the Cepheid V340 Nor is centrally located in the cluster and has the same reddening as the cluster so is assumed to be a member. The Cepheid QZ Nor lies at a distance of two cluster radii out from the cluster centre (Walker 1985b) but is still assumed to be a cluster member in the list given in Feast & Walker (1987). QZ Nor may also be an overtone pulsator (Walker 1985b).

vdBergh1 Our photometry is tested against the photoelectric data of Arp (1960) in Table 8.3. There is no significant offset between the two data sets and no evidence of any scale dependent error. Turner et al. (1998) has compared their photometry to that of Arp (1960) and finds good agreement.

The cluster vdBergh1 was only given a short exposure of 300s in U and 90s in B and V. The U-B:B-V diagram is therefore not very well populated at faint U magnitudes. There is a fair amount of scatter in the U-B:B-V diagram for this cluster, the average value is E(B-V)=0.9±0.09. This is larger than the value of previous estimates. Turner et al. (1998) obtained a minimum value of E(B-V)=0.66, this clearly fits the the edge of the B-type stars. However, for the 10 stars closest to CV Mon, Turner finds E(B-V)=0.80±0.02. Arp (1960) obtained 0.76 (no error) but the spread in the U-B:B-V diagram is such that the larger value would also have been acceptable. 0.77 was used for the cluster reddening by Laney & Stobie (1994).

The distance modulus obtained here is μ=14.10±0.19. This is larger than previous results. Arp (1960) found μ_o=10.94 assuming $R = 3$ which implies μ=13.28 and Turner et al. (1998) found μ_o=11.08±0.03 assuming $R = 3.2$ which implies a distant modulus of μ=13.64. The difference between our value and other values is mostly due to the increased measurement of the reddening. The value used by Laney & Stobie (1994) was 13.6.

The membership of CV Mon to the cluster has been determined by Turner et al. (1998) using radial velocity measurements, evolutionary arguments and by it's location in the cluster.

Trumpler 35 There is little photoelectric data available for this cluster. We compare our photometry against the photoelectric observations of Hoag et al. (1961) and find reasonable agreement in the B and V-bands. The agreement in the U-band is slightly less good but as there is only one source of comparison and as the JKT photometry appears

to agree well for other clusters we suggest that the U-band data photometry is accurate.

The data in the U-B:B-V diagram for the cluster Trumpler 35 (TR35) shows quite a large spread. This is probably caused by differential reddening. Rather than trying to correct for this, a mean value of E(B-V)=1.19±0.07 is taken for the reddening. Shown on the TR35 panel in Figure 8.15 are two lines. The dashed shows the value of E(B-V)=1.03 taken from Turner (1980) and the solid line shows the value of 1.19 adopted here. Using this, the distance modulus to the cluster is then estimated as 15.00±0.18 by fitting to the centre of the data. This is slightly different to the value of Turner (1980) who obtained μ_o=11.6±0.16 with R = 3, which implies μ=14.70, as Turner fitted more to the edge of the B-V:V diagram rather than the centre. The values for the reddening and distance modulus used by Laney & Stobie (1994) (E(B-V)=0.92, μ=14.37) are lower than those of Turner (1980) and those obtained here.

The membership of RU Sct to Trumpler 35 using the reddening and evolutionary status of the Cepheid is discussed and supported by Turner (1980). However RU Sct does lie 15′ away from TR35 (Turner 1980).

NGC6823 was only observed in non-photometric conditions. We use the photoelectric data of Guetter (1991) for calibration purposes. Guetter (1991) has compared his photometry with that of Hiltner (1956) and Hoag (1961) and finds that there are only small offsets of less than 0.04mags in U-B between the different data sets. Turner (1979a) finds good agreement with the photometry of Hiltner (1956) which agrees well with the photometry of Guetter (1991) used for calibration here.

NGC6823 suffers from differential reddening (Turner 1979a), perhaps to an even greater extent than Trumpler 35. Again shown in the panel for the cluster NGC6823 in Figure 8.15 are two lines. One is for E(B-V)=0.53 taken from Feast & Walker (1987) and the other is E(B-V)=0.85 which fits the centre of the B-type stars. The value of E(B-V)=0.85±0.07 is assumed here.

The distance modulus with this reddening is then 13.52±0.20. Turner (1979a) found μ_o=11.81 but there is no mention of the value of R or the reddening so a value of μ cannot be inferred. Laney & Stobie (1994) assume a reddening of E(B-V)=0.44 and μ=14.08, which is slightly larger than the value found here.

The Cepheid SV Vul is only thought to be associated with the cluster NGC6823 (Feast & Walker 1987) as it lies a few arcmins away from the cluster centre.

NGC129 was only observed in non-photometric conditions so previous work had to be relied upon for the calibration, as discussed in detail in Section 8.2.4. The reddening

for this cluster is E(B-V)=0.57±0.05. This value is slightly larger than previous values, Turner, Forbes & Pedreros (1992) obtained 0.47 but they fitted to the least reddened edge of the O and B stars rather than to the centre of these stars. Arp et al. (1959) found 0.53 (with no error) for the reddening which is in agreement with the value found here.

The distance modulus obtained assuming a value of 0.57 for the reddening is μ= 12.67±0.14. This value is similar to the value of μ=12.60 inferred from the values of μ_o=11.11 and R=3.2 obtained by Turner, Forbes & Pedreros (1992). Arp et al. (1959) find μ_o=11.0±0.15 for the distance modulus of the cluster, which with R = 3 implies μ=12.59, again in agreement with the value found here. DL Cas is not included in the study by Laney & Stobie (1994) due to the Northerly latitude of the cluster NGC129.

Kraft (1958) found a measurement of -14±3 km s^{-1} for the radial velocity of the cluster NGC129. He found that the Cepheid itself had a radial velocity of -11kms^{-1}. Given that the error on any individual measurement is estimated to be around 1.5km s^{-1}, DL Cas is assumed to be a member of NGC129.

NGC7790 The U-B:B-V diagram for this cluster appears quite clean and well defined. However, Figure 8.15 shows that NGC7790 has a U-B:B-V diagram where the data poorly fits the ZAMS line. We show one ZAMS shifted to fit the OB stars which implies E(B-V)= 0.59 and another shifted to fit the F stars which would imply E(B-V)=0.43. Most previous estimates are closer to that for the OB stars. This poor fit of the ZAMS to the U-B:B-V data in the case of this cluster is particularly significant since it contains three Cepheids (see Table 8.2).

The U-B:B-V diagram for NGC7790 has a history of controversy. The original UBV photoelectric photometry of Sandage (1958) of 33 11< V<15 stars was criticised by Pedreros et al (1984). A check of 16 stars with the KPNO CCD seemed to confirm that Sandage's U-B and B-V colours were too blue by +0.075 and +0.025 mag respectively, although few details were given of errors etc. However we find good agreement between the work here and the photometry of Sandage, see Table 8.3, Fig. 8.19 and Fig 8.20. A direct comparison of the photoelectric observations of Sandage to the photographic observations in Pedreros (see Table 1 in Pedreros et al. 1984) implies that the differences in the colours of Sandage are too blue by $\sim$ 0.04 mags for both the U-B and B-V colours and when we compare our CCD photometry to 22 of the brightest photographically observed stars from Pedreros, we find similar colour differences of around 0.04 mags in both the U-B and B-V colours.

The U-B:B-V diagram in Pedreros et al. (1984) seemed to give the same sort of ill

fitting ZAMS throughout the range $0.3<B\text{-}V<1.2$ as found in this work. The suggestion was that the problem might lie in Sandage's photometry which Pedreros et al had used for calibration. However, we have tested various zero points for the $U\text{-}B$ and $B\text{-}V$ colours. With our own zero points for the colours we find the ill fitting ZAMS and even if we apply the corrections suggested by Pedreros to the offsets found between our photometry and the photographic data of Pedreros we still find an ill fitting ZAMS.

Romeo et al (1984) used CCD data to obtain $BVRI$ photometry for this cluster to V=20. In the absence of U, their only route to E($B\text{-}V$) was via fitting the shape of the $B\text{-}V:V$ and $V\text{-}I:V$ CMD and they obtained E($B\text{-}V$)=0.54±0.04. They used the Sandage (1958) photometry for calibration in B and V and checked against the B,V photoelectric photometry of 10 stars by Christian et al (1985). They tested the faint photographic photometry of Pedreros et al and found scale errors at $B>17$ and $V>15$ in the sense that Pedreros et al were too bright. In $B\text{-}V$, however they claimed better agreement with $B\text{-}V_{\text{Pedreros}}$ being ~0.1mag too red. A comparison of our photometry and that of Romeo suggests that there is a scale error for $B>17$ and $V>16$ but the extent of this is less than in the comparison of the Pedreros et al. data with the Romeo data, at V=17 our data is brighter than Romeo's by 0.1 mag whereas the Pedreros et al. data is brighter by 0.2 mags and similar differences are found in the B data. We conclude that the photometry in this study agrees well with the photometry of Sandage (1958) and is better agreement with the CCD data of Romeo et al. (1989) than that of Pedreros, (see Fig. 8.22 for more details).

We have also attempted to check to see if contamination is causing the ill fitting ZAMS. Figure 8.6 shows the $B\text{-}V:V$ and the $U\text{-}B:B\text{-}V$ diagrams for the cluster NGC7790. The $B\text{-}V:V$ diagram has been trimmed so that only the stars that lie very close to the ZAMS remain. As these stars have the correct combination of distance and reddening to lie almost on the main sequence then it is likely that they are main sequence stars. The same stars are then used to produce the $U\text{-}B:B\text{-}V$ colour-colour diagram. The same UV deficit around the F-type stars that appears in Figure 8.15, when all the stars in the field are used, appears in Figure 8.6. Figure 8.7 compares the colour-colour diagram found from this work with the work of Sandage (filled triangles), Pedreros (filled squares) and photometry obtained by Fry and Carney (open triangles, B Carney, private communications). The Sandage points do not go deep enough to test the shape of the data but the Pedreros et al. and Fry and Carney points do and the poor match of the data to the ZAMS is seen. Therefore, we believe that the ill fitting ZAMS to the $U\text{-}B:B\text{-}V$ diagram is not caused by

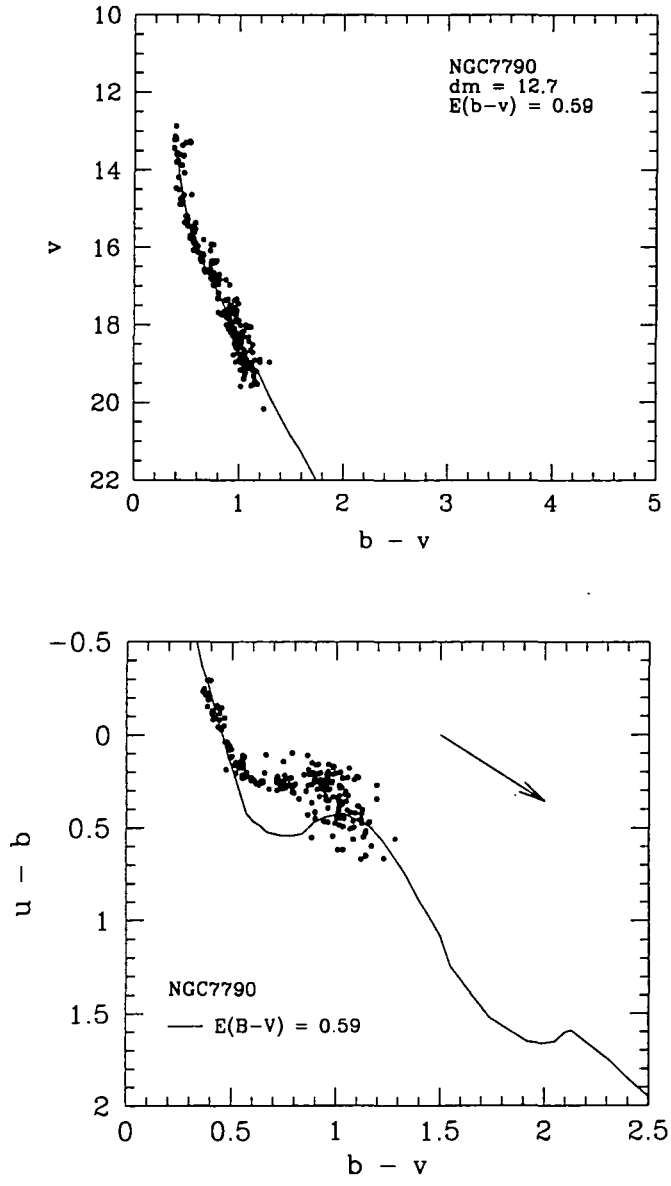

Figure 8.6: The top panel shows the *B-V:V* colour-magnitude diagram for the cluster NGC7790 but with only the stars that fit the zero age main sequence shown. The lower panel shows the same stars as they appear on the *U-B:B-V* diagram. The U-V deficit is still clearly apparent among the stars that are most likely to be members of the cluster.

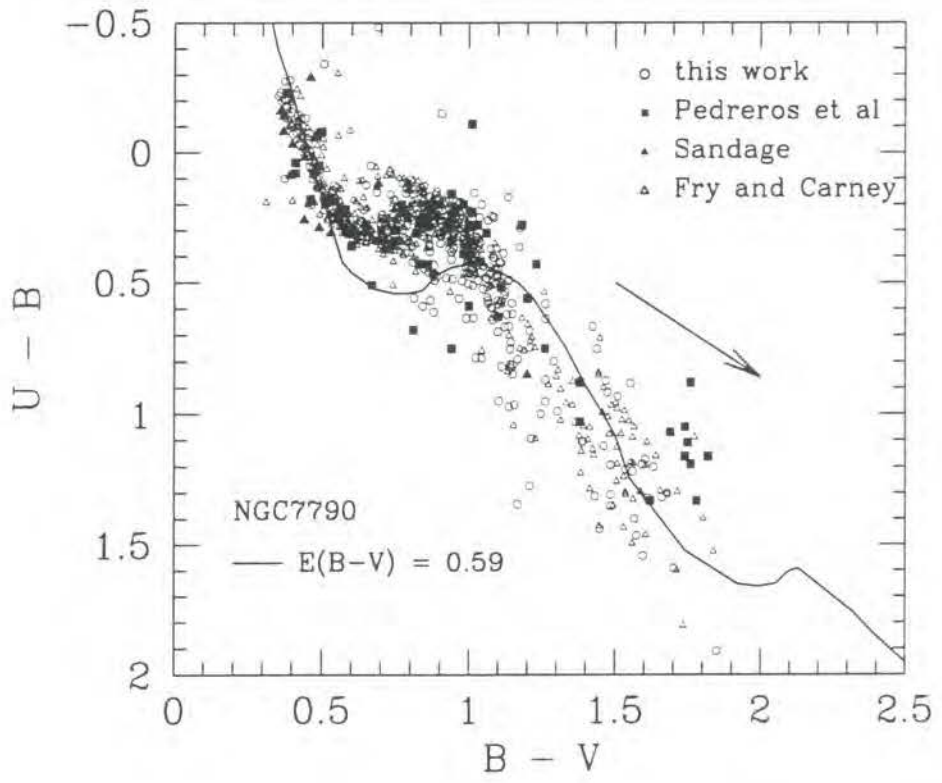

Figure 8.7: The open circles show the photometry of this work. The solid triangles show the photometry of Sandage which does not go deep enough to test the ill fitting ZAMS. The results of Pedreros et al. (1984)(filled squares) and Fry and Carney (open triangles) show the same effect as the photometry in this work and confirms the poor fit of the ZAMS *U-B:B-V* relation in NGC7790

contamination from foreground or background stars or by errors in the photometry but is a real feature in the data.

Thus our estimate of the NGC7790 reddening based on the *U-B:B-V* colours of OB stars, E(*B-V*)=0.59±0.05, is between the E(*B-V*)=0.52 ±0.04 of Sandage (1958) and the E(*B-V*)= 0.63±0.05 of Pedreros et al(1984) who used similar techniques.

Assuming the estimate of E(*B-V*)=0.59±0.05 from the OB stars, the distance modulus which fits *V:B-V* is then μ=14.37±0.13, which is close to the value of 14.35, inferred from the value of μ_o=12.8±0.15 with $R = 3$ found by Sandage (1958). Romeo et al. (1989) assume a formula for the value of R, which gives R=3.09 with E(*B-V*)=0.54 and (*B-V*)$_o$=0 and therefore give a distance modulus of $\mu \sim$14.3, as they measure μ_o=12.65±0.15. Pedreros et al. (1984) assume $R = 3.28$ and therefore a value of μ =14.40 is inferred from their value of μ_o=12.3.

Sandage (1958) states that the membership of CF Cas, CEa Cas and CEb Cas to NGC7790 is almost certain due to the position of the Cepheids on the CMD.

8.4 P-L Relation

Using the values for the distance modulus and the reddening towards the cluster, we proceed to determine the P-L relation. As well as the reddening and the distance modulus of the cluster, the apparent magnitude of each of the Cepheids is required. Where possible these come from Laney and Stobie (1993, 1994) who have high quality *V* and *K*-band measurements for most of the Cepheids in this study. The clusters NGC7790 and NGC129 lie at northerly latitudes so are unobservable from SAAO and so there are no magnitudes from Laney and Stobie for these Cepheids. The *K*-band data for DL Cas and CF Cas comes therefore from Welch et al. (1985) and the *V*-band data and the periods are taken from Feast & Walker (1987).

The reddening obtained from the ZAMS fitting is that of the cluster OB stars. Schmidt-Kaler (1982) found that when the effect of the colour difference between the OB stars and the Cepheid is taken into account,

$$E(B-V)_{ceph} = E(B-V)_{clus}[0.98 - 0.09(< B_o > - < V_o >)_{ceph}] \qquad (8.2)$$

gives a good approximation to the reddening of the Cepheid. The values for $(< B_o > - < V_o >)_{ceph}$ come from Feast & Walker (1987).

Figure 8.8 shows the M_V-M_K - Log(P) relation for the Galactic Cepheids. This is

Cepheid	Log(P)	$E(B\text{-}V)_{ceph}$	<V>	<K>	M_V	M_K	M_V^{LS}	M_K^{LS}	$\Delta(M_V)$	$\Delta(M_K)$
V367	0.799	1.27	11.604	6.662	-3.854	-5.035	-3.755	-4.992	-0.099	-0.043
U Sgr	0.829	0.45	6.685	3.952	-3.861	-5.266	-3.781	-5.214	-0.080	-0.052
EV Sct	0.490	0.61	10.131	7.028	-2.641	-3.947	-2.191	-3.553	-0.450	-0.394
WZ Sgr	1.339	0.50	8.023	4.565	-4.587	-6.747	-4.657	-6.789	0.070	0.042
TW Nor	1.033	1.24	11.670	6.319	-3.756	-5.314	-3.797	-5.411	0.041	0.097
QZ Nor	0.730	0.26♯	8.866	6.622	-3.177	-4.637	-3.128	-4.588	-0.049	-0.049
V340 Nor	1.053	0.33	8.375	5.586	-3.906	-5.695	-3.797	-5.640	-0.109	-0.055
CV Mon	0.731	0.74	10.306	6.576	-3.794	-5.302	-3.305	-4.896	-0.489	-0.406
RU Sct	1.294	0.93♯	9.465	5.071	-4.979	-6.568	-5.186	-6.775	0.207	0.207
SV Vul	1.654	0.44♯	7.243	3.920	-5.150	-7.126	-6.028	-8.004	0.878	0.878
DL Cas	0.903	0.52	8.970	5.920	-3.679	-5.069	-3.860	-5.360†	0.181	0.291
CF Cas	0.688	0.55	11.140*	7.990	-3.201	-5.019	-3.170*	-4.850†	-0.031	-0.169
CEa Cas	0.711	0.55	10.920*	NA	-3.592	NA	-3.390*	NA	-0.031	NA
CEb Cas	0.661	0.55	10.990*	NA	-3.524	NA	-3.330*	NA	-0.031	NA

Table 8.5: Comparison between the work here and that of Laney & Stobie (1994). * indicates where values are taken from Feast & Walker (1987) and † where the values are inferred from Welch et al. (1985). ♯ indicates where the Cepheid reddenings have been corrected (see text)

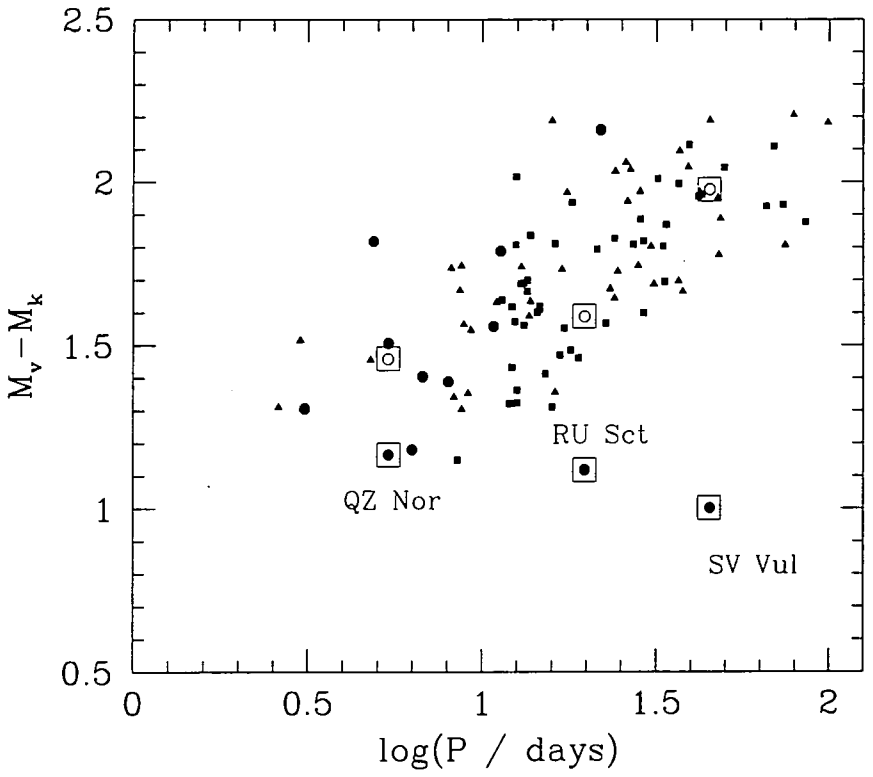

Figure 8.8: The Galactic Cepheids are shown by the solid circles, the LMC Cepheids by the triangles and the SMC Cepheids by the squares, taken from Laney & Stobie (1994). The three solid circles in a box are QZ Nor, RU Sct and SV Vul. The reddenings of these Cepheids are corrected to the values given in Laney & Stobie (1993), indicated here by the boxed open circles.

calculated via

$$M_v - M_k = V - K - 0.91\mathrm{E}(B-V)_{\mathrm{ceph}} \tag{8.3}$$

with the Cepheid reddenings obtained from the cluster reddenings using equation 8.2. The triangles show the same relation for the LMC Cepheids and the squares are for the SMC Cepheids. These are taken from Tables 2 and 3 in Laney & Stobie (1994). Three of the Cepheids, (from left to right in Figure 8.8) QZ Nor, RU Sct and SV Vul seem to have the wrong M_v-M_k colours for their periods. The clusters containing RU Sct and SV Vul suffer both from the presence of differential reddening and from the fact that the Cepheids lie at some distance away from the cluster. Figure 8.8 indicates that the reddening local to RU Sct and SV Vul may be somewhat different from the average value of the cluster reddening. We correct for this using the space reddenings given in Laney & Stobie (1993) which are more local to the Cepheids (see for example Turner 1980). Note that because there is difficulty obtaining the Cepheids true reddening from the cluster reddening for these two Cepheids, we do not include them in the best sample (see later in this Section).

The Cepheid QZ Nor lies just on the edge of the period colour relation. Walker (1985b) notes that the Cepheid QZ Nor lies away from the centre of the cluster NGC6067, at a distance of two cluster radii so again the reddening of the cluster may not be appropriate for the reddening of the Cepheid. Laney & Stobie (1993) take the value for the Cepheid reddening from Coulson & Caldwell (1984) of E(B-V) = 0.265, derived from BVI_c reddenings. This is the reddening used to calculate the position of the open circle in Figure 8.8 and which we assume for QZ Nor henceforth. However, the effect of changing the reddening of the Cepheid to E(B-V) = 0.265 changes the distance modulus to the LMC by only a small amount as the Cepheid lies in a well populated part of the P-L relation. This Cepheid is not included in the best sample though (again see later in this Section).

To determine the absolute magnitude of the Cepheid, the apparent magnitude has to be corrected for reddening and distance. First of all the extinction coefficient is required. We follow Laney & Stobie (1994) and use

$$R = 3.07 + 0.28(B-V)_o + 0.04\mathrm{E}(B-V) \tag{8.4}$$

where the E(B-V) here is the Cepheid value to take into account the effect of Cepheid colour on the ratio of total to selective extinction. The Cepheid reddening comes from equation 8.2 except for the three corrected values. Then the reddening free magnitudes of the Cepheids are

$$V_o = V - R(\mathrm{ceph})\mathrm{E}(B-V)_{\mathrm{ceph}}$$

$$K_o = V_o - V + K + \frac{R(\text{ceph})E(B-V)_{\text{ceph}}}{1.1} \quad (8.5)$$

The expression for K_o has the form given above as the extinction coefficient in the K-band is approximately one tenth of that in the V-band. To obtain finally the absolute magnitude in each waveband, the dereddened distance modulus, $\mu_o = \mu - R(\text{clus})E(B-V)_{\text{clus}}$, has to be subtracted off V_o and K_o

The P-L relation can now be determined. We consider two samples, one where we consider all the Cepheids available to us and another where the Cepheids RU Sct, SV Vul and QZ Nor are removed due to the problem of obtaining the correct reddening and the question of cluster membership for SV Vul. The zero points are obtained by fixing the slope and obtaining the least squares solution using the Galactic Cepheids in this study. These are summarised in Table 8.6. The slopes that are considered are the slopes from Laney & Stobie (1994) which are the best fitting slopes to all the Cepheid data (Galactic open cluster Cepheids, LMC and SMC Cepheids) in their study. The slopes are -2.874 in the V-band and -3.443 in the K-band. Also considered is -2.81 in the V-band as this is the slope of the LMC Cepheids and the slope used by Feast & Catchpole (1997).

Once the slope and zero point of the PL relation is fixed, the distance modulus to the LMC can be calculated. We estimate this by finding the shift between LMC and galactic Cepheids, assuming a common slope and zero point. Laney & Stobie (1994) give the period and the dereddened V and K-band magnitudes, V_o and K_o of 45 LMC Cepheids. The distance to the LMC is then given by

$$< \mu_o >= \frac{1}{N}\sum_{i}^{N}(m_o - (\delta * \log(P) + \rho)) \quad (8.6)$$

where m_o represents the dereddened apparent magnitude in each waveband and δ and ρ are the values for the slope and zero point as given in Table 8.6. Our PL(V) and PL(K) relations are shown in Figures 8.9, 8.10 and 8.11. The zero points and distance to the LMC used in Figures 8.9, 8.10 and 8.11 come from the best sample, as given in Table 8.6.

Taking the value for the PL(V) zeropoint for the best sample with the -2.874 slope used by Laney & Stobie (1994) gives ρ=-1.205±0.11 which is in good agreement with the value of ρ=-1.197±0.06 found by these authors and implies a distance modulus of 18.50±0.12 for the LMC in agreement with 18.50±0.09 found by Laney & Stobie.

Our PL(V) zeropoint is slightly lower than the zeropoint and LMC distance obtained from an analysis of Hipparcos trigonometrical parallaxes of nearby Galactic Cepheids by Feast & Catchpole (1997). They obtained ρ=-1.43±0.1 for the Galactic PL(V) zeropoint

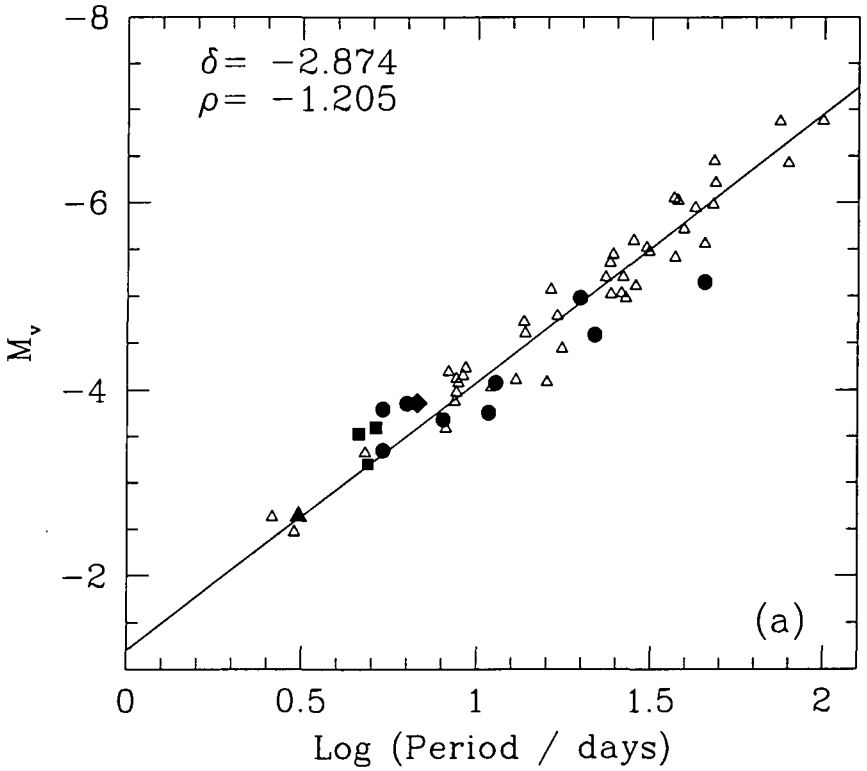

Figure 8.9: The *V*-band Cepheid P-L relation with a slope of -2.874. The solid symbols are the Galactic Cepheids, marked specifically are U Sgr (diamond), EV Sct (filled triangle), CF Cas, CEa Cas and CEb Cas (squares). The zeropoint and slope is shown in the Figure. The open triangles show the P-L relation for the LMC Cepheids. The periods and magnitudes are taken from Table 2 and 3 in Laney & Stobie (1994) and the distance modulus to the LMC is the 'best' values from Table 8.6 in this work.

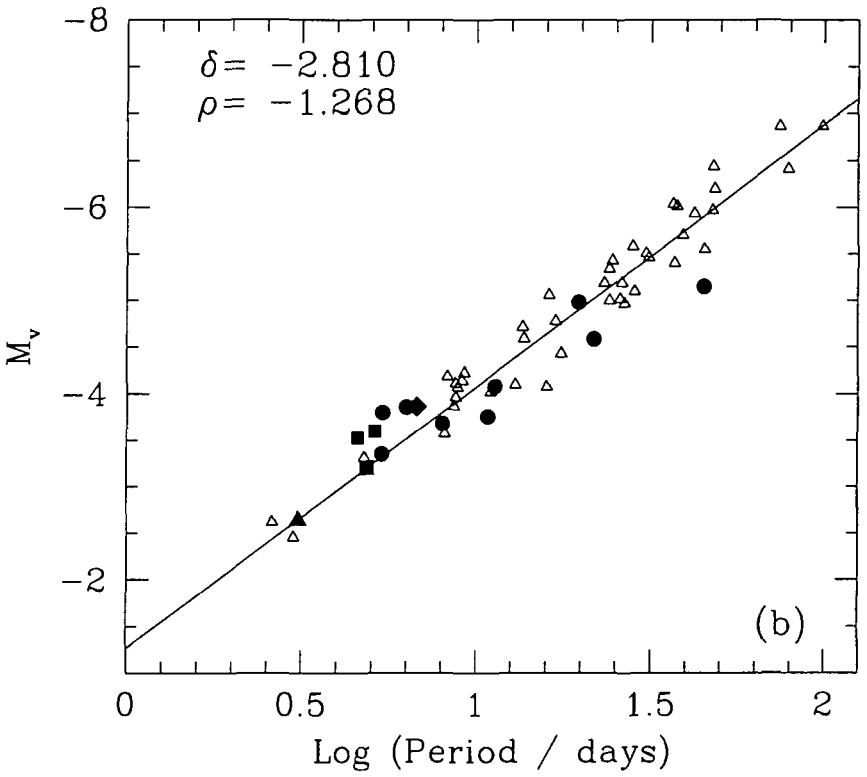

Figure 8.10: The *V*-band Cepheid P-L relation with a slope of -2.81. The symbols are the same as for Figure 8.9. The zeropoint and slope is shown in the Figure. The periods and magnitudes are taken from Table 2 and 3 in Laney & Stobie (1994) and the distance modulus to the LMC is the 'best' values from Table 8.6 in this work.

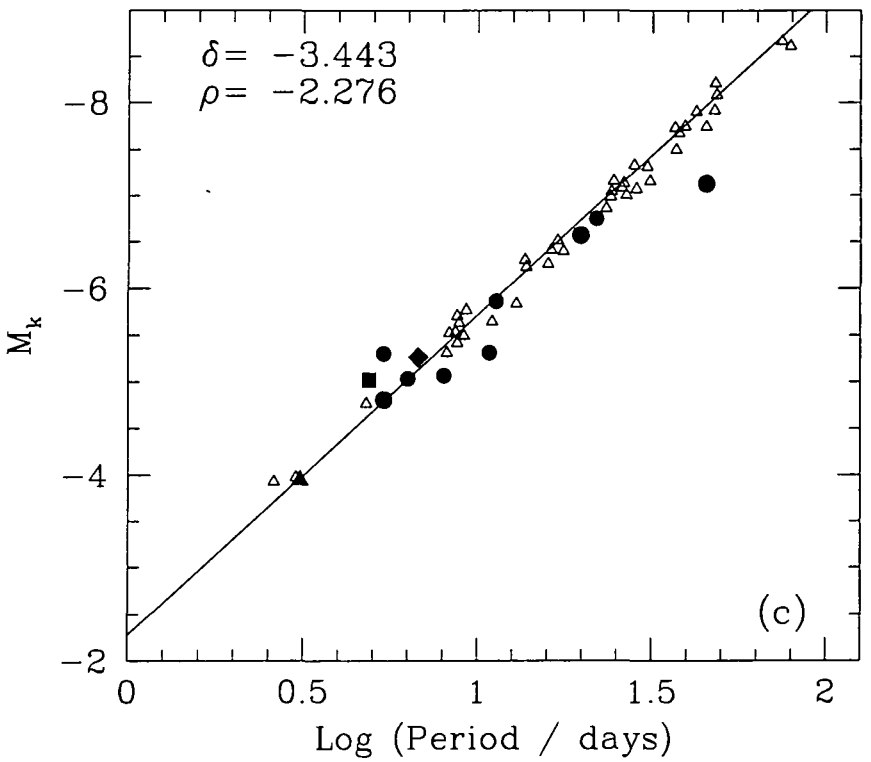

Figure 8.11: The K-band Cepheid P-L relation with a slope of -3.443. The symbols are the same as for Figure 8.9 but there is no K-band data for the Cepheids CEa Cas and CEb Cas. The zeropoint and slope is shown in the Figure. The periods and magnitudes are taken from Table 2 and 3 in Laney & Stobie (1994) and the distance modulus to the LMC is the 'best' values from Table 8.6 in this work.

Band	Sample	Slope(δ)	Zeropoint(ρ)	μ_o(LMC)
V	All(14)	-2.874	-1.209±0.100	18.50±0.036
V	All(14)	-2.810	-1.268±0.095	18.48±0.036
K	All(12)	-3.443	-2.193±0.105	18.46±0.019
V	Best(9)	-2.874	-1.205±0.110	18.50±0.036
V	Best(9)	-2.810	-1.268±0.106	18.47±0.036
K	Best(9)	-3.443	-2.276±0.106	18.54±0.019
V	logP<1.0(9)	-2.874	-1.410±0.072	18.71±0.036
V	logP<1.0(9)	-2.810	-1.456±0.072	18.67±0.036
K	logP<1.0(7)	-3.443	-2.378±0.103	18.65±0.019

Table 8.6: The zeropoints for the P-L relation and distance modulus to the LMC

for an assumed slope of δ=-2.81. This can be compared to ρ=-1.268±0.11 obtained for our best sample with the same slope. They used the same 45 Laney & Stobie (1994) Cepheids as used here to obtain a metallicity corrected LMC distance modulus μ_o=18.70±0.10. We note that their semi-theoretical metallicity correction to the LMC Cepheid V magnitudes increases the distance to the LMC, which is in the opposite sense to most empirically determined estimates of the effects of metallicity on Cepheids (eg Kennicutt et al 1998). Subtracting their metallicity correction leads to an LMC distance modulus μ_o=18.66±0.10 which can be directly compared with our best value of μ_o=18.47±0.11, combining the errors on the slope and distance modulus to the LMC from Table 8.6. Our PL(V) estimates for the LMC distance modulus are similar to those of Laney & Stobie and below those of Feast & Catchpole.

The K-band P-L relation is tighter for the LMC Cepheids and for the Galactic Cepheids, see Figure 8.11, and the slopes are closer with the LMC Cepheids giving δ=-3.27±0.04 and the best sample of Galactic Cepheids giving δ=-2.81±0.12. We note in passing that the slope of the Galactic Cepheid PL(K) relation is now flatter than the δ=-3.79±0.1 slope found in the Galactic Cepheid sample of Laney & Stobie. Assuming the -3.443 slope used by Laney & Stobie our best sample in Table 8.6 gives a PL(K) zero-point of ρ=-2.276±0.11 which implies an LMC distance of μ_o=18.54±0.12 which remains in good agreement with the value μ_o=18.56±0.06 found by Laney & Stobie. (1994). The smaller error of Laney & Stobie is due to their larger numbers of calibrators although it must be said that many of their extra calibrators (8/12) are in associations rather than clusters and frequently given half-weight in P-L fits. Moreover, they have not included

the 4 cluster Cepheids in NGC129 and NGC7790. Therefore we believe that our result supercedes the Laney & Stobie result with our bigger error estimate perhaps being a more realistic indication of the actual errors.

Finally, as our overall estimate of the LMC distance, we take the average of the PL(V) and PL(K) estimates in the best sample of Table 8.6 which gives μ_o=18.51±0.10. We conclude that although in the case of individual clusters we have markedly improved the distance and reddening estimates, our new estimates of the zeropoint of the PL relation and thus the distance to the LMC are close to previous values.

8.5 Discussion

We now discuss the most intriguing new result in this study, which is that the Solar metallicity ZAMS may not always fit the *U-B:B-V* data in individual clusters. This is not the first time an effect like this has been seen. Turner (1993b) saw poorly fitting *U-B:B-V* ZAMS for the open cluster Roslund 3. Turner interpreted this as evidence for the young, B-type stars having a cocoon of circumstellar dust around them. This cocoon of dust then increases the reddening of the B-type stars as compared to the F and G type stars, causing the ill-fitting *U-B:B-V* ZAMS. The O and B type stars in the *U-B:B-V* diagram of Roslund 3 show a large spread around the ZAMS as the amount of excess dust would probably vary from star to star. Excess reddening could perhaps also be so strong that it was causing some O and B stars to be so reddened that they appeared as F type stars. However, the O and B type stars in the *U-B:B-V* diagram for NGC7790 are very tight so the shape of the *U-B:B-V* diagram is unlikely to be caused by excess dust.

The next possibility we consider is that the effect might be due to stellar evolution, However, the CMD for the clusters look unevolved even at AOV as might be expected for clusters which have Cepheid variables which are therefore expected to be less than 10^8 years old.

We also consider whether the discrepancy between the main sequence fitted distance and the Hipparcos parallax to the Pleiades could explain our result. If it were assumed that all open clusters had roughly the same composition then the different forms for *U-B:B-V* that we find might be taken as evidence that the colours of main sequence stars may not be unique. This possibility has also been discussed as an explanation of the problem with the MS fitted distance to the Pleiades (van Leeuwen 1999) and if it proves relevant in that case it will certainly also be worthy of further consideration here.

If the reddening vector in *U-B:B-V* varied as a function of Galactic position then this would also affect our results. However, at least in the case of NGC7790 it seems that for whatever relative shift in *U-B* and *B-V*, the ZAMS still has the wrong shape to fit the observed colour-colour relation.

U-band CCDs have improved greatly over the last decade but still there are questions as to the suitability of *U*-band CCDs for this type of work. However, the cluster NGC7790 has been observed by three different groups, Pedreros et al. (1984), Fry and Carney (although the photometry has not been published) and this study. The CCD photometry from each group was obtained on two different telescopes (Kitt Peak and JKT) and Pedreros et al. (1984) obtained further photographic data on the Canada-France Hawaii 3.6m telescope. Shown in Figure 8.7 is a comparison between these three sets of photometry and the same discrepancy is seen. It is perhaps hard to imagine that the problem with the *U*-band CCD on the JKT is exactly the same as the problems with the *U*-band CCD on the Kitt Peak telescope. Also if the *U*-band CCD is causing spurious features in the colour-colour diagram of NGC7790, why does the colour colour diagram of M25 and NGC129 appear to match the ZAMS line so well?

The final possibility is that metallicity is affecting the F stars' *U-B* colours in some of these clusters. Qualitatively there is some evidence supporting this suggestion. First, 'line blanketing' is well known to redden the *U-B* colours of metal rich stars at F and G and low metallicity sub-dwarfs are known to show UV excess as the reverse of this case (e.g. Cameron (1984)). The open circles in Figure 8.12 shows the colour-colour diagram for NGC7790. The solid line is the solar ZAMS and the dashed line shows the colour-colour empirical relation for stars aged 10^8 years with Fe/H=-1.7 (Alex Vazdekis, private communication). The dashed line shows some level of agreement with the bulk of the stars in the colour-colour diagram with *B-V* colours of ~0.9. For larger *B-V* colours the stars do not agree with the high metallicity line though. Some of these stars could be foreground or background contamination stars.

Second, there is some suggestion that the cluster, NGC7790, that shows a UV excess, lies outside the solar radius while NGC6664 which is redder in *U-B* at F lies inside the solar radius (see Figure 8.13). M25 which lies closest to the Sun also fits the solar metallicity *U-B:B-V* diagram as well as any of the clusters. Given that metallicity in the Galaxy is expected to decrease with Galactocentric radius, this is suggestive of a metallicity explanation. Many of the other clusters' *U-B:B-V* diagrams are either too noisy due to differential reddening (Trumpler35, NGC6823) or too obscured to reach the F stars

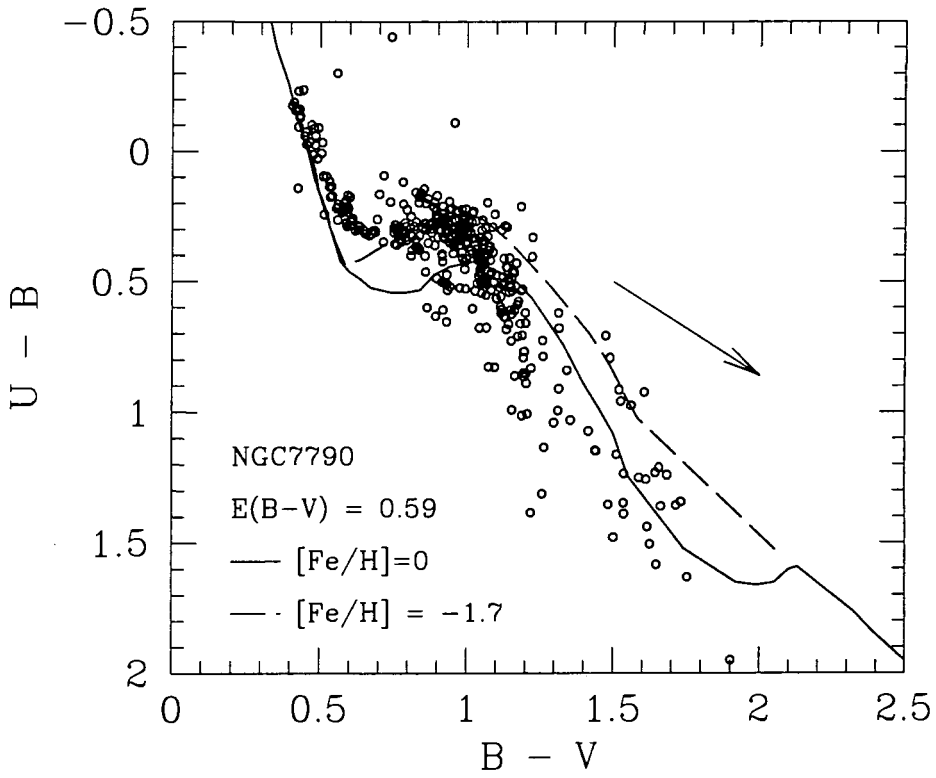

Figure 8.12: The open circles show the colour-colour diagram of NGC7790. The solid line shows the solar ZAMS and the dashed lines shows the empirical relation for stars with age 10^8 years and a metallicity of [Fe/H]=-1.7 reddened to E(B-V)=0.59.

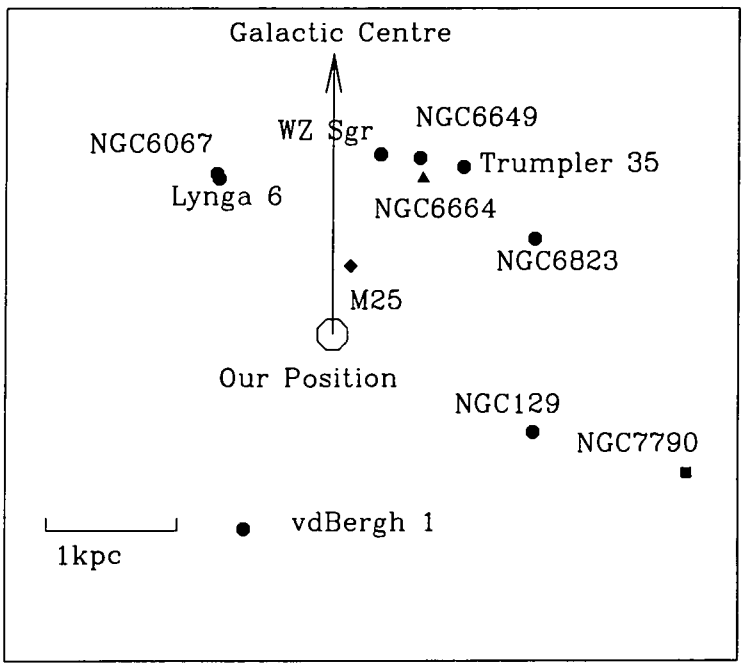

Figure 8.13: The approximate positions of all the clusters considered here. The clusters NGC6664 (triangle), M25 (diamond) and NGC7790 (square) are highlighted. These symbols are the same as in Figure 8.9

(NGC6649, Lynga 6, vdBergh1) to further test this hypothesis. However, NGC6067 forms a counter-example to any simple gradient explanation, since it seems to have a normal UBV plot and lies inside the solar position This would have to be accommodated by allowing a substantial variation on top of any average metallicity gradient.

However, quantitatively the case for metallicity is less clear. The size of the UV excess seen is much larger in the case of NGC7790 than expected on the basis of previous metallicity estimates of these clusters, or of any measurement of the amplitude of the Galactic metallicity gradient. Using the Fe/H vs Δ *U-B* relations of Carney (1979) or Cameron (1985) it would be concluded that NGC7790 showed Δ *U-B* $\sim$ 0.2mag which corresponds to Fe/H$\sim$-1.5. Thus clusters which on the basis of their Main Sequences and the presence of Cepheids, must be less than 10^8 yr old, would be implied to have near halo metallicity. Previously, Panagia & Tossi (1991) find Fe/H$\sim$-0.3 for these 2 clusters. Also Fry & Carney (1997) find Fe/H=-0.2$\pm$0.02 for NGC7790 while finding Fe/H=-0.37$\pm$0.03 for NGC6664 based on spectroscopy of the Cepheids in these clusters themselves. Also according to the Galactocentric metallicity gradient which is usually taken to lie in the range -0.02-0.1dex kpc^{-1} (Rana 1991), there should only be on the average Δ Fe/H $\sim$ 0.3 in the range of metallicity covering these clusters.

On the other hand, it should be noted that Panagia & Tosi's Fe/H estimates are based on more poorly measured estimates of UV excess than those presented here and also that there is little agreement between the metallicity estimates of Fry and Carney and those of Panagia & Tosi. Measuring the metallicity of the Cepheids themselves as attempted by Fry and Carney is difficult since the effective temperature is a function of the light curve phase and a small difference in estimated temperature can make a large difference in metallicity. Also the Galactocentric metallicity gradient at least as measured for open clusters depends on relatively poor *U-B* photometry at the limit of previous data from Janes (1979), Cameron (1985) and Panagia & Tossi (1991). In any case, it is well accepted that the dispersion in metallicity around the mean gradient is indeed high, with the range -0.6<Fe/H<+0.3 at the Solar position. Furthermore Geisler (1987) using Washington photometry to estimate metallicity also found an example of a cluster, NGC 2112, only $\sim$ 0.8kpc outside the solar radius with Fe/H=-1.2, although this cluster is older than those discussed here.

However, it would also seem that the tightness of the P-L relations in Figures 8.9, 8.10 and 8.11 could form a final argument against the idea that NGC7790 has Fe/H$\sim$-1.5. If the metallicity of the cluster NGC7790 was really Fe/H$\sim$-1.5 then at given B-V,

main sequence stars would be sub-dwarfs with ~1mag fainter absolute V magnitudes than normal solar metallicity main sequence stars (see Cameron, 1984, Figures 5,6). Thus since we have used a normal Main Sequence to derive the distance to NGC7790, it is surprising that the Cepheids in these clusters lie so tight on the P-L relation when they should be a magnitude too bright if the low metallicity hypothesis is correct. The only way that the low metallicity hypothesis for NGC7790 could survive this argument is if it were postulated that the effect of metallicity on the Main sequence star magnitude and the Cepheid magnitude were the same - then the effect of our derived distance modulus being ~1 magnitude too high would be cancelled out by the fact that the Cepheid is actually sub-luminous by 1 magnitude because of metallicity which would leave the Cepheid tight on the P-L relation as observed. This might not be too contrived if a low metallicity Cepheid prefers to oscillate about its subdwarf, rather than solar metallicity, zero-age luminosity (at fixed effective temperature) position on the Main Sequence. This would lead to a strong implied metallicity effect on the Cepheid PL(V) and PL(K) zeropoints; the implication would be that $\frac{\delta M}{\delta Fe/H}$ ~0.66 in the sense that lower metallicity Cepheids are fainter. This coefficient is within the range that has been discussed for the empirical effects of metallicity on Cepheids by Kennicutt et al. (1998) and Gould (1994) although the most recent work by Kennicutt et al. (1998) appears to give a lower value of $\frac{\delta M}{\delta Fe/H}$ ~0.24±0.16 again in the same sense.

The immediate effect on the distance to the LMC with Fe/H=-0.3 is that our estimate of its distance modulus would decrease from 18.5 to 18.3. However, since all that is determined at the LMC is the slope of the P-L relation, then the zeropoints we have derived in Table 8.6 from the Galactic Cepheids would still refer the P-L relation to the Galactic zeropoint. The ultimate effect on H_o would then be decided by the metallicity of the Cepheids in the galaxies observed by, for example, the HST for the Distance Scale Key project Ferrarese et al. (2000), by Tanvir et al. (1995) in the case of the Leo I Group and by Saha et al. (1999) in the case of SNIa. Zaritsky (1994) have measured metallicities for these galaxies already but if the dispersion in Cepheid metallicity is as large as it is implied to be in the Galaxy then there may be some signature in a wider dispersion in the Cepheid P-L relations in at least the high metallicity cases. The possibility of detecting this signature is currently being investigated (Allen and Shanks, 2000 in preparation)

Obviously the most direct route to checking the metallicity explanation for the anomalous behaviour seen in the U-B:B-V diagrams is to obtain medium-high dispersion spectroscopy for a sample of F stars in NGC7790 and NGC6664 to determine the metallicity

directly for these main sequence stars. Currently proposals are in to use WHT ISIS spectrograph for this purpose.

8.6 Conclusions

We have presented colour-colour diagrams and colour-magnitude diagrams of a sample of galactic clusters which contain or are associated with Cepheids. All the clusters have been observed using similar methods and the data reduction and extraction has also been done with similar techniques. The use of the improved *U*-band data has allowed powerful new checks of previous E(*B*-*V*) estimates over a wide range of magnitudes. In order to estimate the reddenings and distance moduli, we have fitted all the clusters in the same manner and have not attempted to correct for differential reddening but have instead taken a simpler approach and fitted the average reddening value of the cluster. In most cases the differences that we have found between values for the reddening and distance modulus are small and where there are significant differences these can mostly be explained by comparing whether the ZAMS fit was made to the centre or to the edge of the colour-colour and colour-magnitude diagrams.

The Cepheid P-L relations found from fitting the best sample are M_V=-2.81×log(P)-1.261 and M_K=-3.44×log(P)-2.276 and a distance modulus to the LMC of 18.47±0.12 in the *V*-band and 18.54±0.11 in the *K*-band giving an overall distance modulus to the LMC of 18.51±0.10, ignoring any possible effects of metallicity. These results for both the PL relations and the LMC distance are consistent with the previous results of Laney & Stobie (1994) although the improved distances and reddenings have increased the errors over what was previously claimed. These increased errors mean that our result for the PL(*V*) relation are almost consistent with the result of Feast & Catchpole (1997) from Hipparcos measurements of Cepheid parallaxes, although this gives rise to an LMC distance modulus of μ_o= 18.66±0.1 as opposed to our μ_o=18.51±0.10.

With the improved *U*-band data, we find that for at least two of the clusters, the data in the *U*-*B*:*B*-*V* two-colour diagram is not well fitted by the solar metallicity ZAMS. One possibility is that significant metallicity variations from cluster to cluster may be affecting the *U*-*B* colours of F- and G-type stars. The problem is that the metallicity variations this would require are much larger than expected for young, open clusters. More work is therefore required to determine the metallicity of the individual main sequence stars in each of the clusters NGC7790 and NGC6664. If metallicity is proven to be the cause of

the anomalous *U-B:B-V* relations, then it would imply that the Cepheid P-L relation in both the visible and the near-infrared is strongly affected by metallicity.

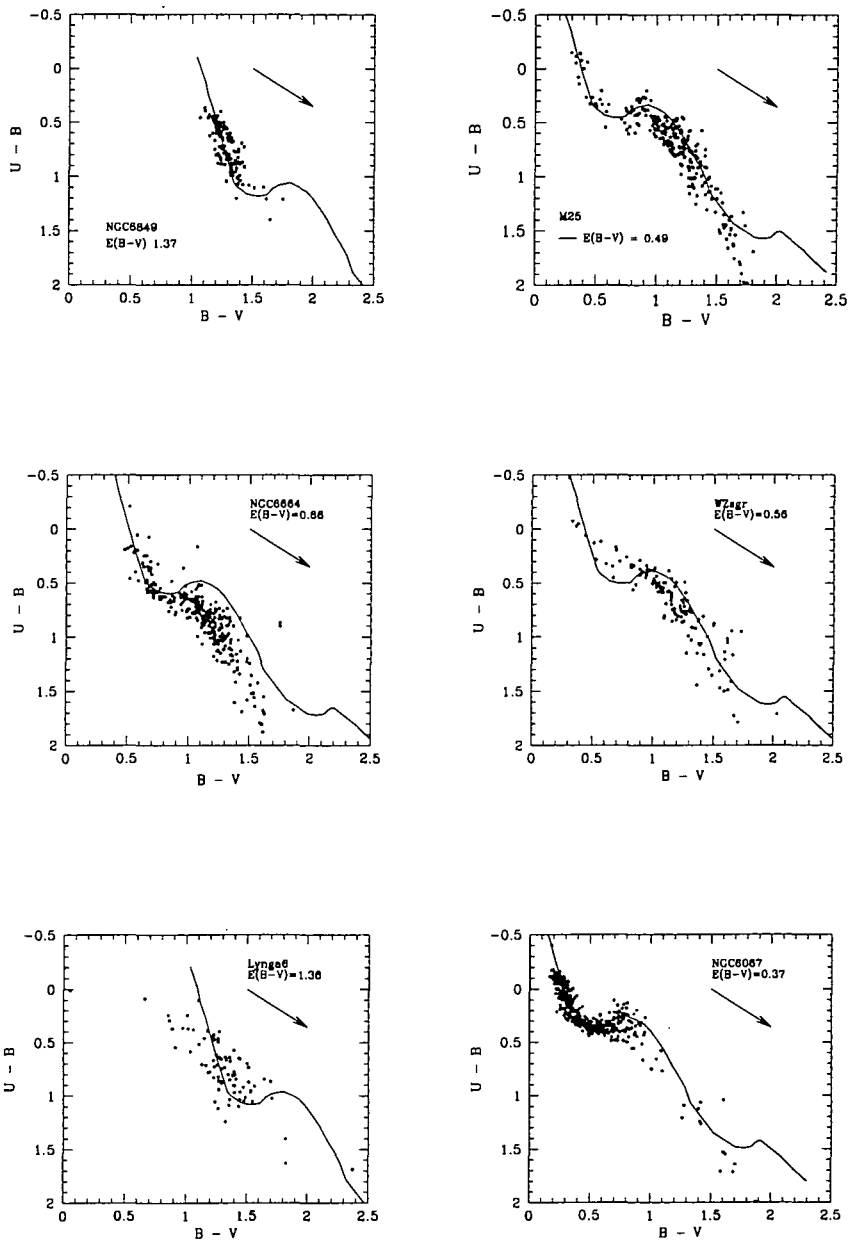

Figure 8.14: The *U-B:B-V* diagrams for the clusters used in the study. The arrow indicates the direction of the reddening vector.

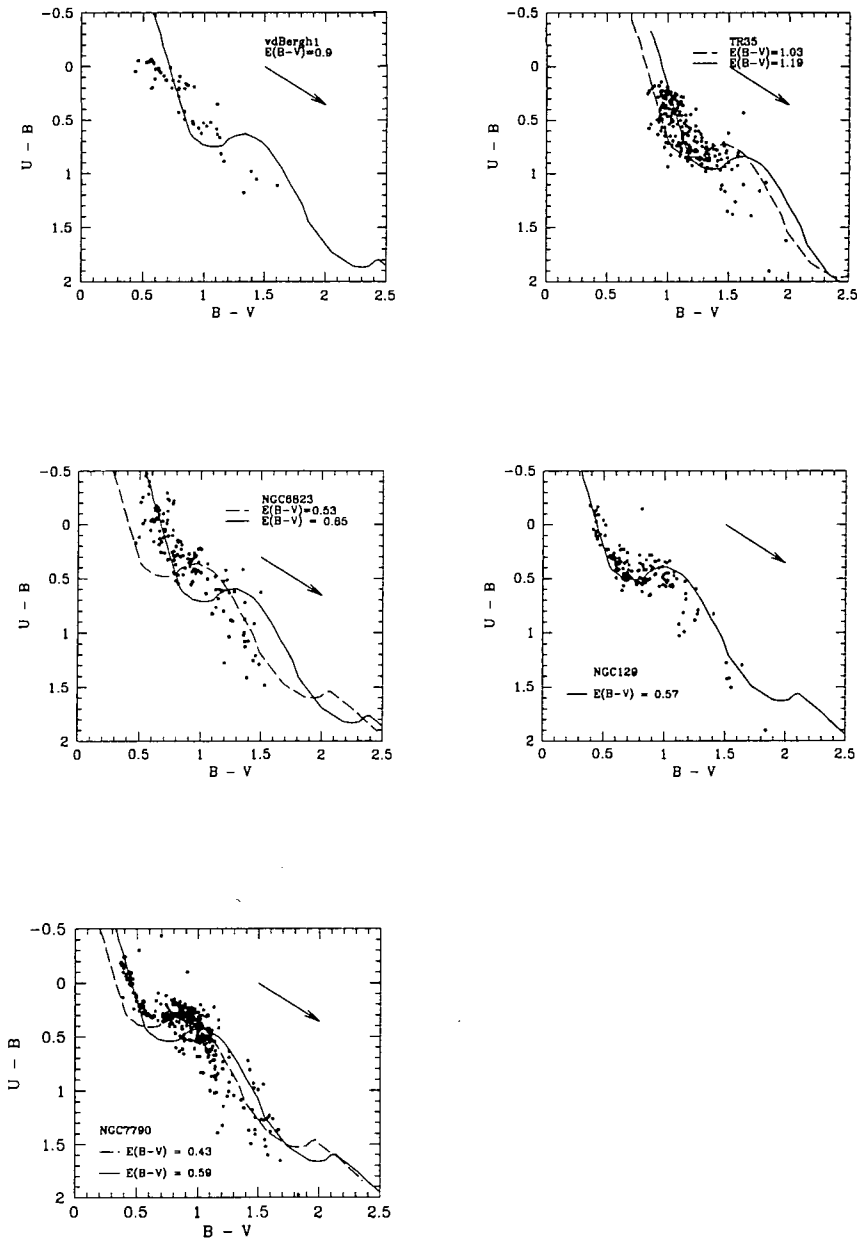

Figure 8.15: More *U-B:B-V* diagrams for the clusters used in the study. The arrow indicates the direction of the reddening vector.

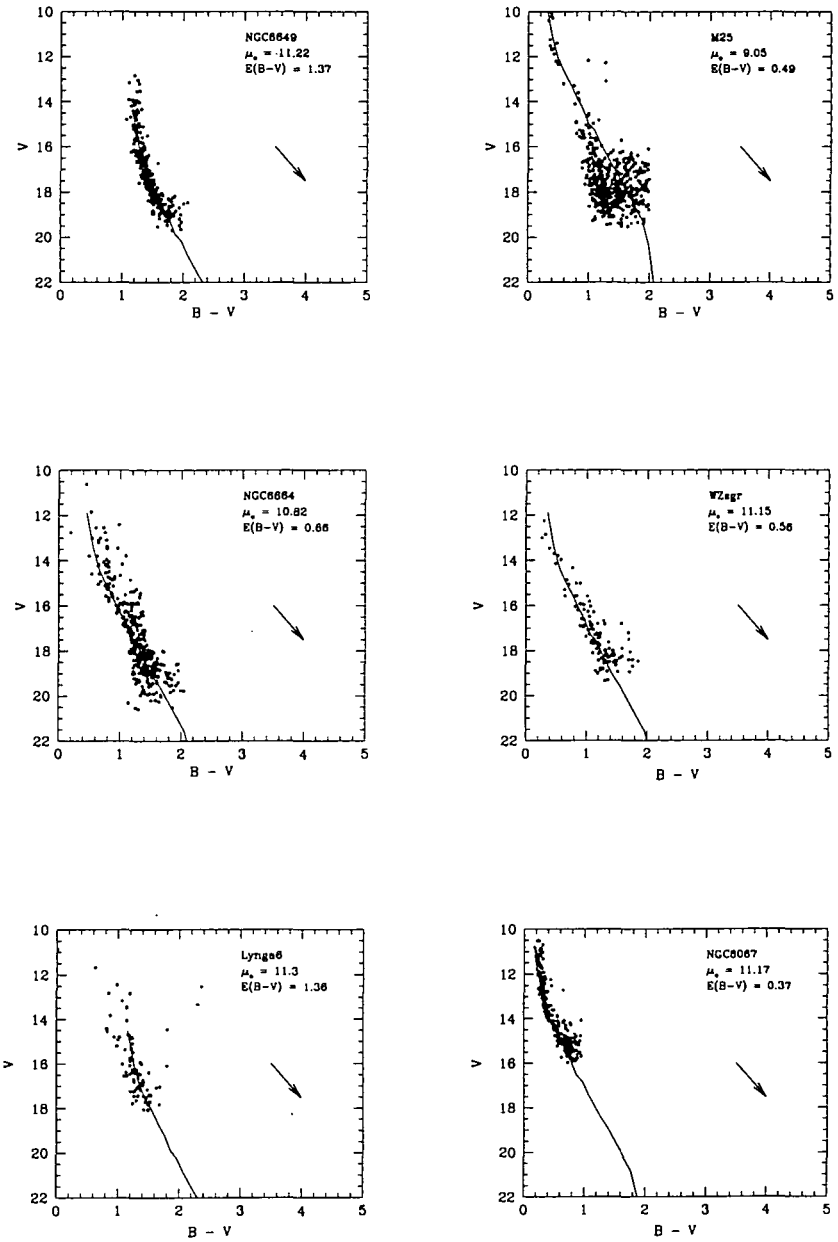

Figure 8.16: The *V-B:V* diagrams for the clusters used in the study. The dereddened distances are shown in the plot assuming $R = 3.1$.

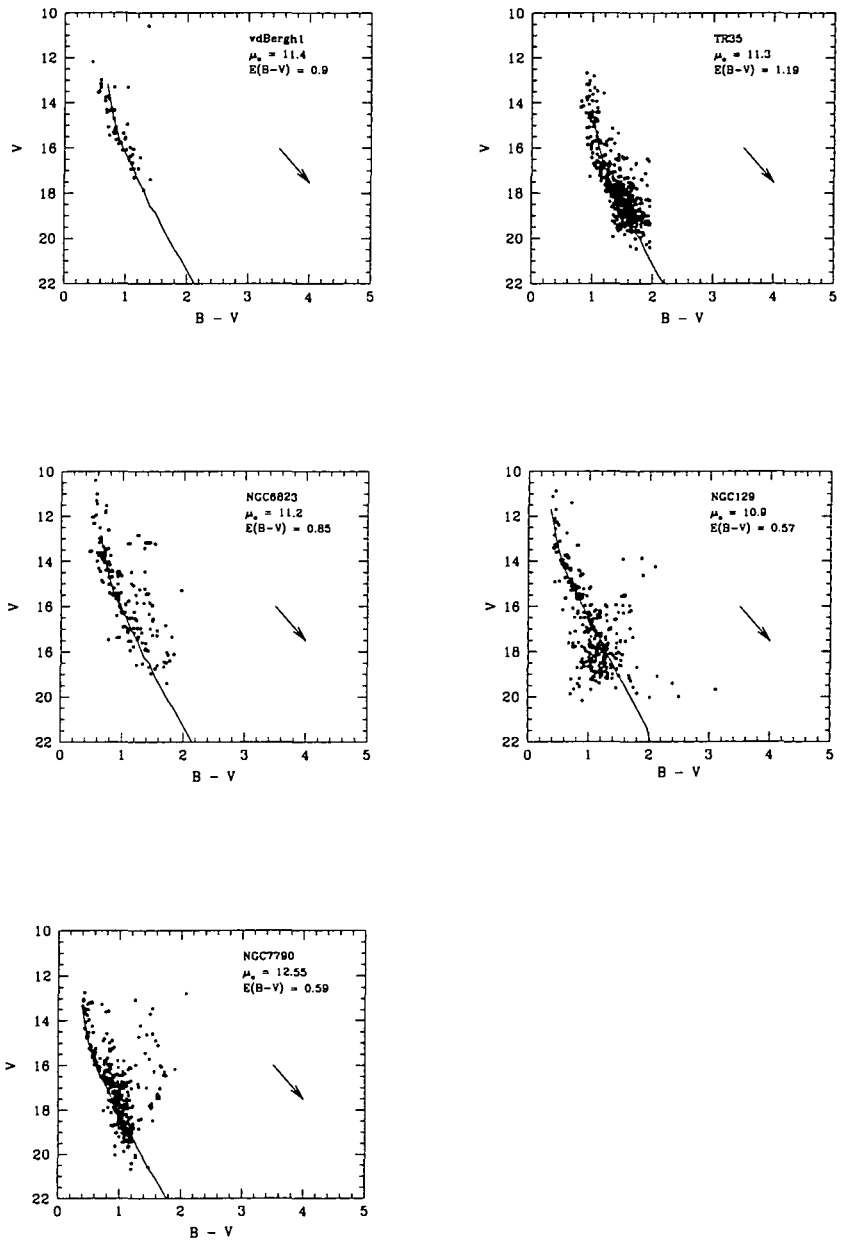

Figure 8.17: Further *V*-*B*:*V* diagrams for the clusters used in the study. The dereddened distances are shown in the plot assuming $R = 3.1$.

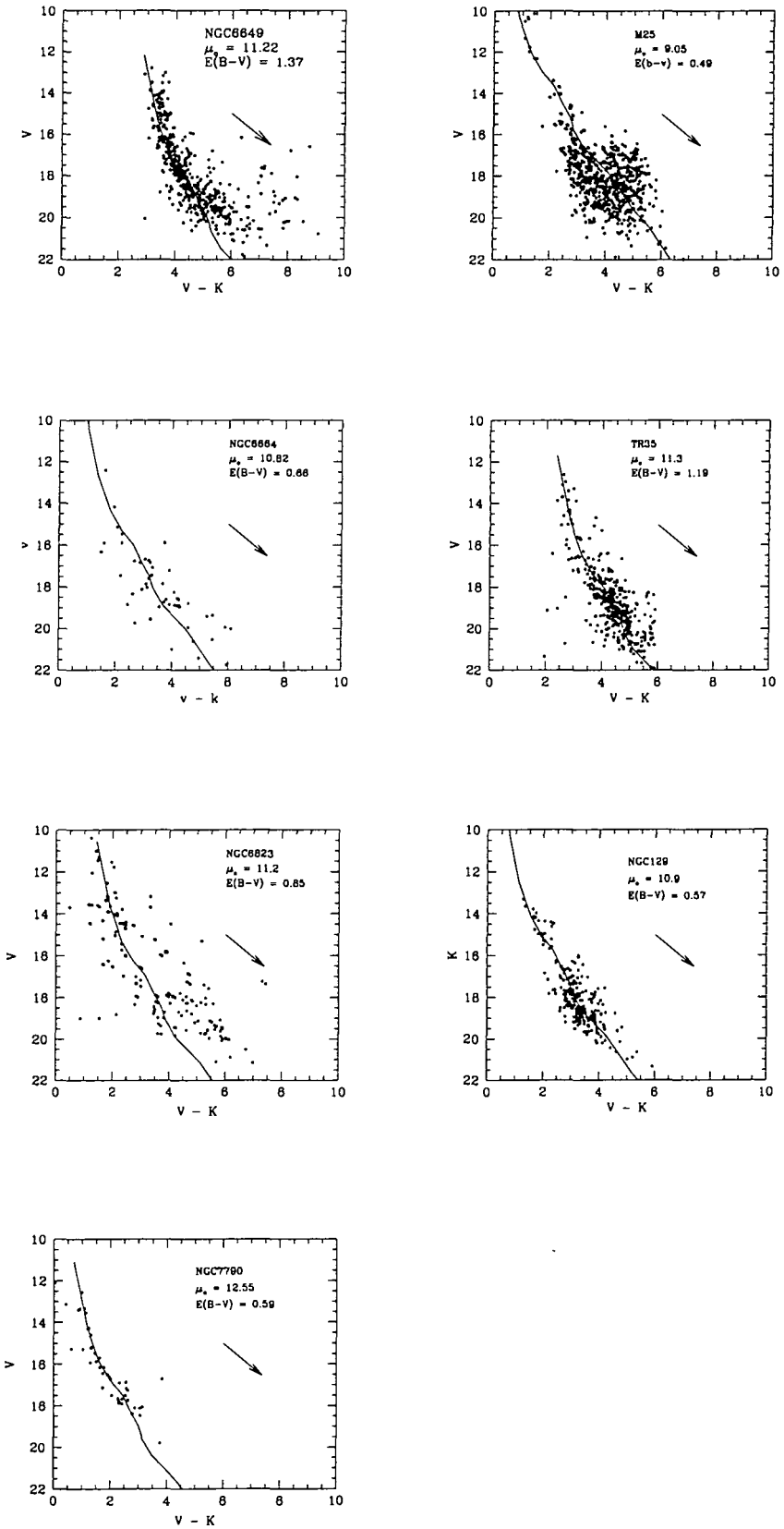

Figure 8.18: The *V-K:V* diagrams for the clusters used in the study. The dereddened distances are shown in the plot assuming $R = 3.1$.

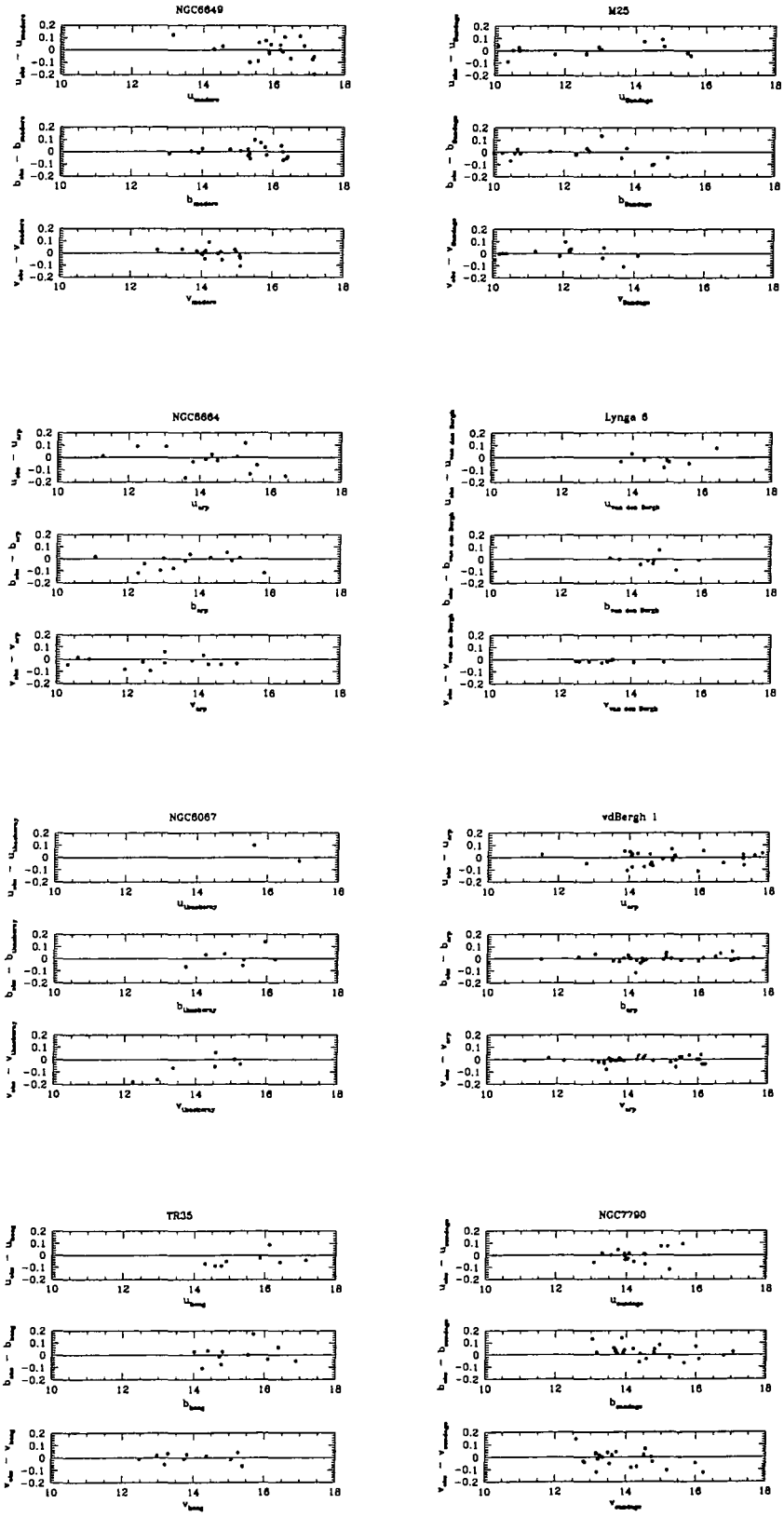

Figure 8.19: A comparison between the zero point compared in this work with the zero point obtained in previous photoelectric studies. The full reference for the comparison is given in Table 8.3. Only the clusters where an independent zero points was obtained are shown.

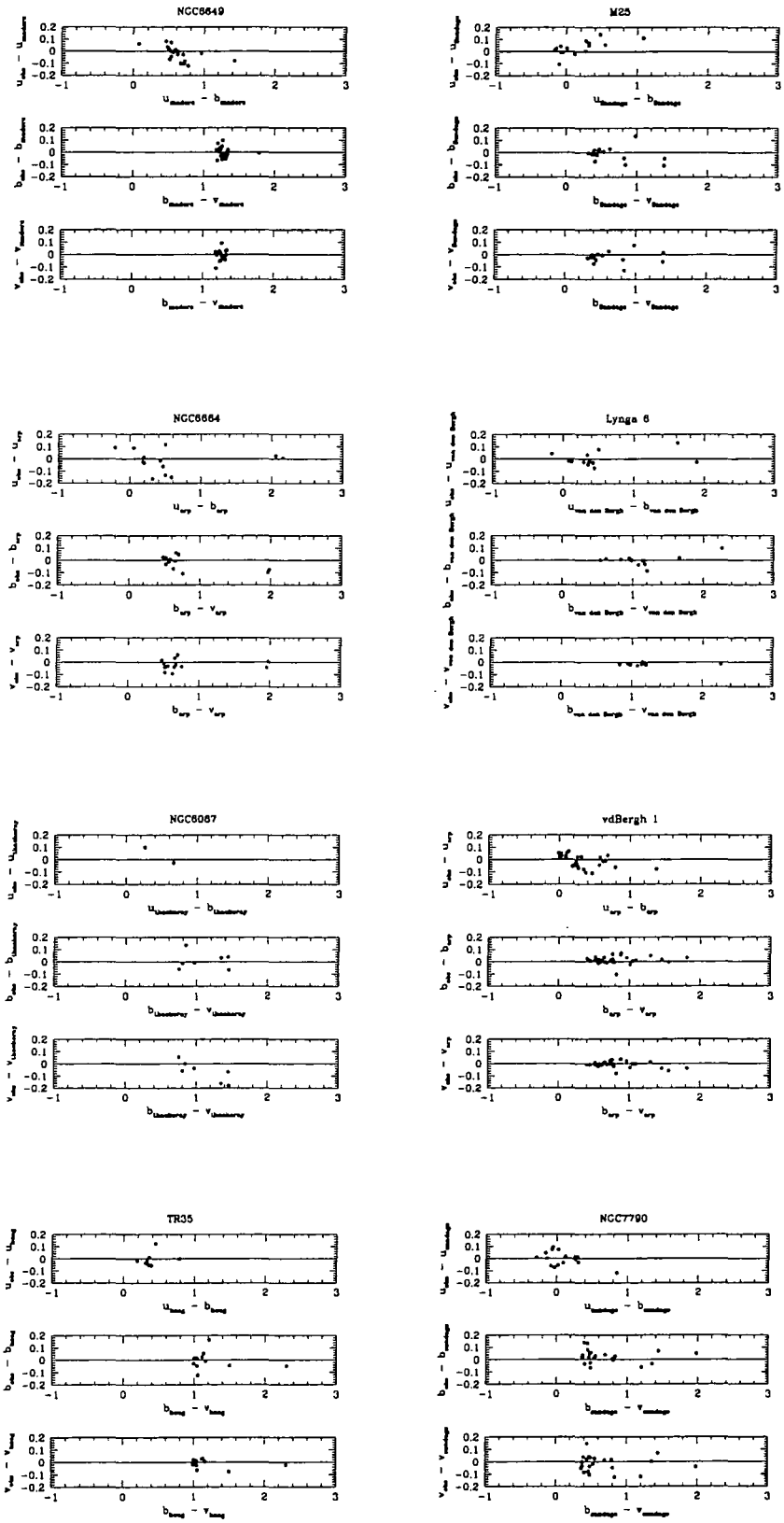

Figure 8.20: A check for any colour dependent relationship between the work in this study and previous photoelectric work. The full reference for the comparison is given in Table 8.3. Only the clusters where an independent zero points was obtained are shown.

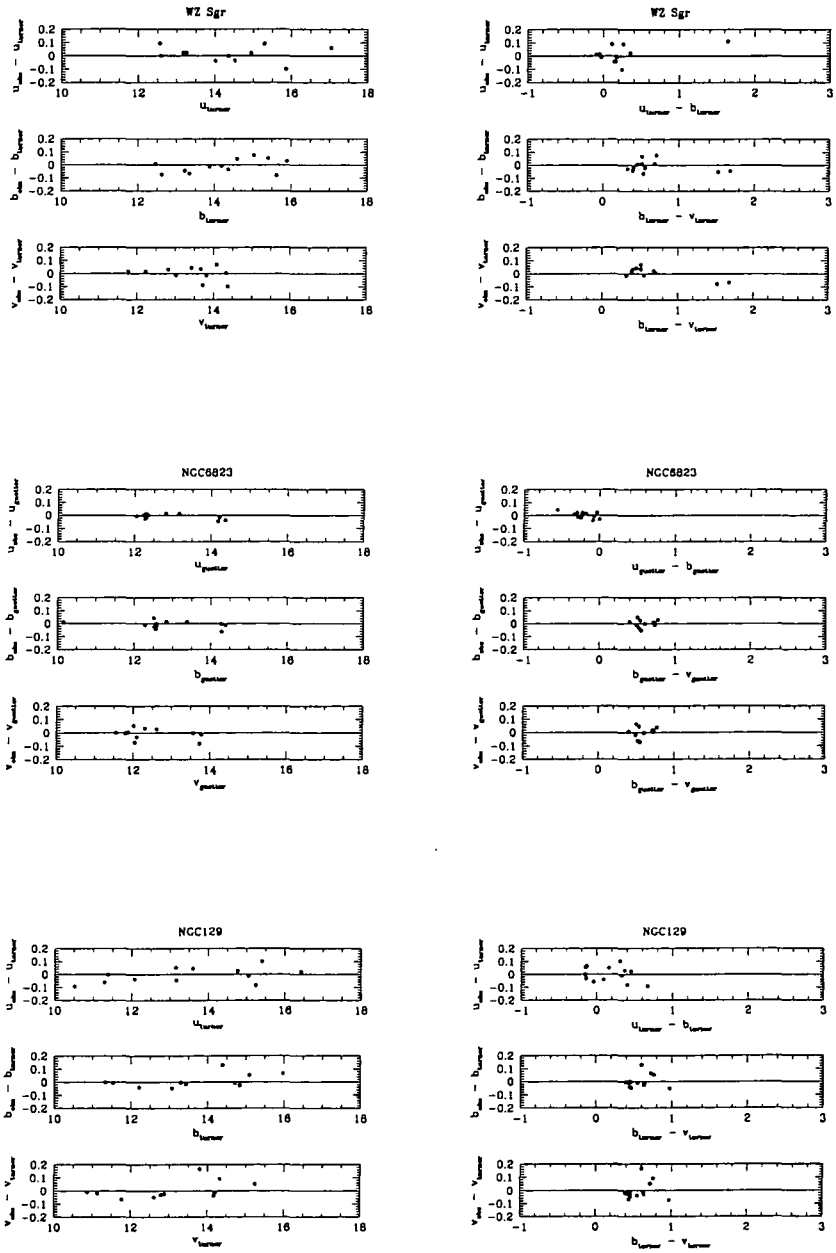

Figure 8.21: The zero points and colour comparison for the clusters where previous work had to be relied upon for calibration purposes.

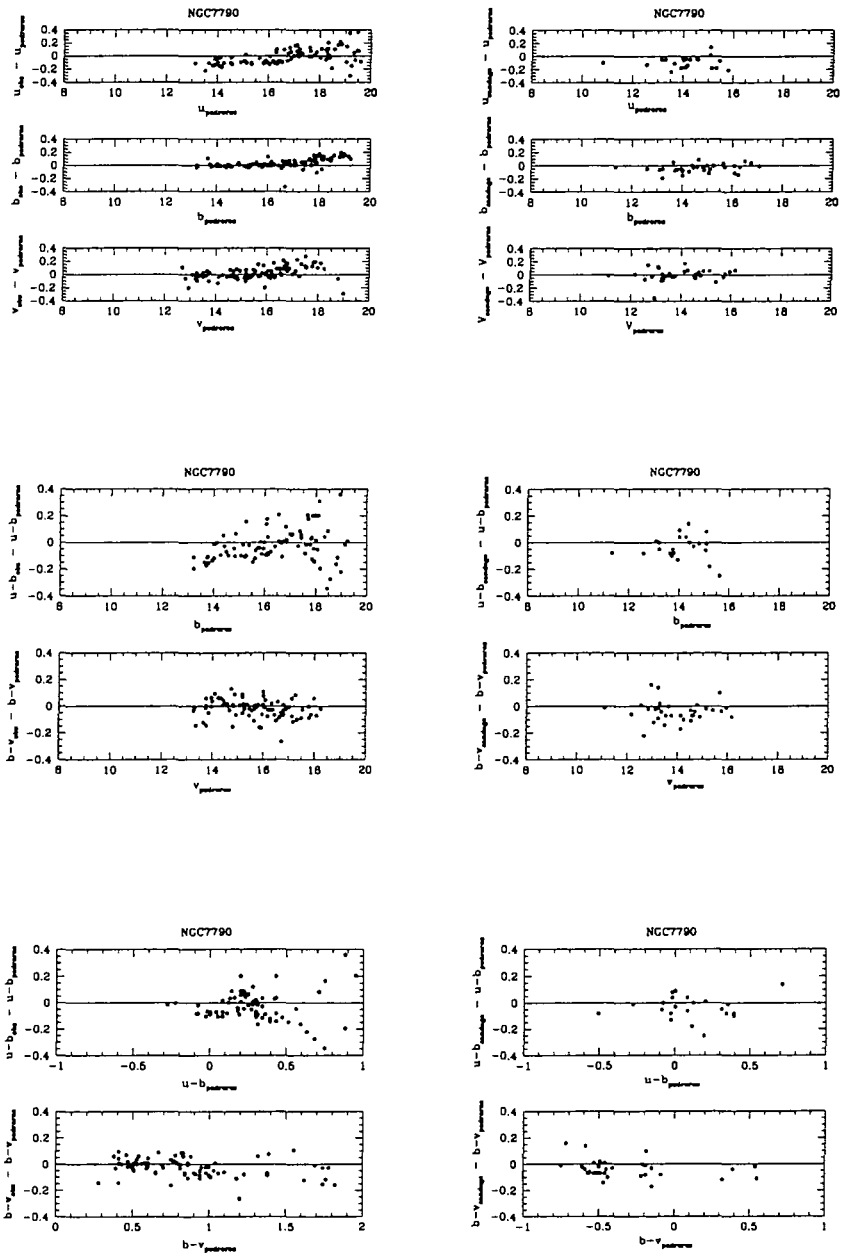

Figure 8.22: Comparison of all the different sources of photometry for NGC7790.

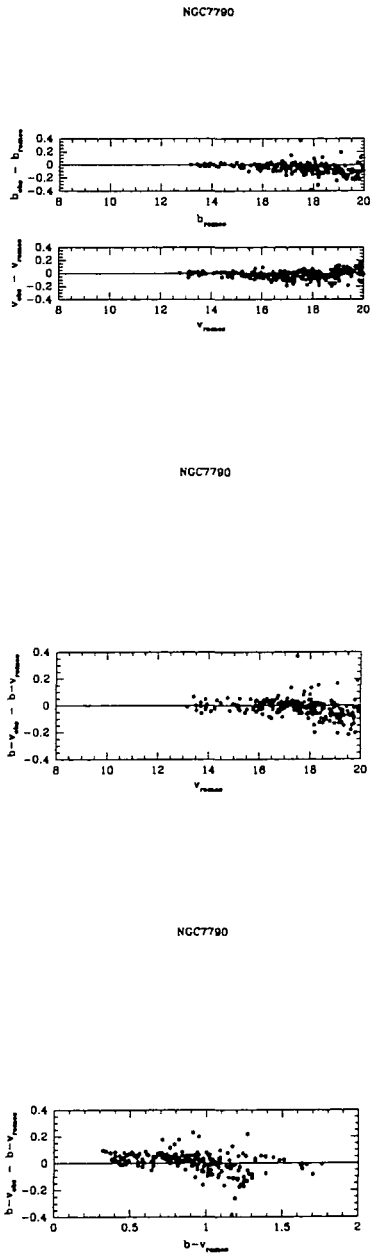

Figure 8.23: Further comparison of all the different sources of photometry for NGC7790. Note there is no U-band data in the study by Romeo et al. (1989)

Chapter 9

Conclusions and Future Work

9.1 Summary of the Main Conclusions

One of the primary aims of this Thesis has been to quantify the large scale structure of the Universe. Structures, such as clusters and filaments of galaxies or QSOs and voids completely empty of objects can be detected simply by looking at the distribution of objects in redshift surveys. However, redshift surveys are only able to discriminate between different models of structure formation if clustering statistics, such as the power spectrum, correlation function and the distribution of counts-in-cells, are measured from the surveys. This is because many of the models of structure formation display clusters, filaments and voids which visually appear quite similar to those seen in the Universe. However, when the structures are quantified, a more robust comparison between statistics from redshift surveys and models of structure formation can be made, allowing some models to be rejected.

We have measured the galaxy power spectrum from the Durham/UKST Survey on scales $10 \lesssim r \lesssim 120h^{-1}$Mpc using volume-limited and flux limited samples. There is good agreement between the power spectra of the optically selected Durham/UKST and Stromlo-APM (Tadros & Efstathiou 1996) surveys but on scales larger than $80h^{-1}$Mpc we obtain more power from the Durham/UKST Survey than is found in the LCRS Survey (Lin et al. 1996). The shape of the power spectrum from the combined, infrared selected, 1.2Jy and QDOT Surveys (Tadros & Efstathiou 1995) is similar to that of the Durham/UKST Survey but it has a lower amplitude and a relative bias squared of $b^2_{\rm rel}$=1.7 is required to match the amplitude of the Durham/UKST survey.

By comparing the shape of the Durham/UKST redshift space power spectrum and the real space power spectrum of the APM survey (Baugh & Efstathiou 1993), we measure the value of the one dimensional velocity dispersion to be 320±140 km s^{-1} and β to

be 0.60±0.35. These values are compared to values obtained from models of large scale structure. We are able to rule out the standard cold dark matter (CDM) model with Ω_m=1 and shape parameter Γ=0.5 with either COBE or cluster normalisation as the velocity dispersion required for these models to match the shape of the Durham/UKST power spectrum is too high. A ΛCDM model has velocities that are too high at the 1-2σ level. The best agreement is with a biased τCDM model with shape parameter Γ=0.2.

Galaxy clustering can be described in more detail if we consider higher order moments of the clustering as well as the two-point statistics such as the power spectrum. We have therefore also applied a counts-in-cells analysis to the Durham/UKST and the Stromlo-APM surveys. We directly measure the skewness, S_3, out to scales of $20h^{-1}$Mpc for the first time. Again we compare to models of structure formation and find that on scales below $20h^{-1}$Mpc a linear bias cannot reconcile S_3 measured from CDM models with S_3 measured from galaxy surveys.

Many of the popular models of structure formation are able to approximately reproduce the clustering at the present day. However, different models evolve differently as we look back in time. The evolution of a model is driven by the cosmological parameters Ω_m and Ω_Λ, so by looking at the clustering of objects at high redshift, such as QSOs, further constraints on models of structure formation and cosmological parameters can be obtained.

The 2dF QSO Survey will contain 25,000 QSOs with redshifts in the range $0.3 \lesssim z \lesssim 3$ and will provide complementary information to that from galaxy redshift surveys. To date, the Survey is not finished but in this Thesis we have made several predictions for the clustering expected in the completed survey using mock 2dF QSO Surveys and we have measured the power spectrum from the QSOs observed prior to January 2000.

We created mock 2dF QSO catalogues using a simple bias scheme suggested by Cole et al. (1998). The correlation function from the mock catalogues has a similar shape and amplitude as that from the January 2000 catalogue. We have tested different estimators of the correlation function and find the results are insensitive to the choice. The errors on the correlation function on intermediate scales, $2.5 < r < 40h^{-1}$Mpc, will be better than 20%, assuming the mock catalogues are reasonable representations of the final survey. The QSO power spectrum will also be measurable and we predict that the power spectrum will be free from distortions due to the window function on scales up to $400h^{-1}$Mpc, assuming Ω_m=0.3 and Ω_Λ=0.7.

Using the first 9,000 QSOs measured in the 2dF QSO survey, we have measured the

power spectrum of optically selected QSOs for the first time. The agreement with the mock catalogues is reasonable over the range of scales $50 < r < 200h^{-1}$Mpc, when the same cosmology as the simulation is adopted. We measure the QSO power spectrum from two bins of redshift and find little evolution in the QSO clustering amplitude. The QSO power spectrum has a similar amplitude to the power spectrum of present day optically selected galaxies but a far lower amplitude than the power spectrum of present day rich clusters (Tadros, Efstathiou & Dalton 1998). By comparing the shape of the QSO power spectrum to power spectra from models of structure formation, we measure the shape, parameterised by Γ, to be ~ 0.1, although incompleteness may be biasing Γ slightly low.

We also apply the Λ test (Alcock & Paczyński 1979), described in Chapter 7, to the mock QSO catalogues. We find that it is not possible to strongly constrain $\Omega_{\rm m}$ from the QSOs as the errors on the clustering measured parallel and perpendicular to the line of sight, $\xi(\sigma, \pi)$, are too large, although β is reasonably well constrained, which puts joint constraints on cosmology and the QSO-mass bias. However, the results from the Λ test can be combined with results from the clustering of galaxies and QSOs to place tighter constraints on cosmological parameters. We predict that if $\Omega_{\rm m} = 0.3$ and $\Omega_\Lambda = 0.7$ is the cosmology of the Universe, then $\Omega_{\rm m} = 0$, Ω_Λ=1 will be rejected with more than 2σ confidence. Results from the Λ test applied to the 2dF QSO Survey will also be independent of results on cosmology from supernova or CMB experiments, providing yet further constraints on the cosmological parameters $\Omega_{\rm m}$ and Ω_Λ

As well as quantifying the structure of the Universe, we have also considered its scale. Using improved *U* band CCD photometry, we have reanalysed 11 Galactic open clusters that contain Cepheids or have Cepheids associated with them. We conclude that the previous estimates of the reddening and distance modulus of the clusters are in reasonable agreement with the new values presented here and thus the previous calibration of the Cepheid P-L relation is very similar to ours. Using our updated values for distance and reddening we measure the distance modulus of the LMC to be 18.51±0.10. However, our deeper photometry allows us to probe further down the ZAMS in the *U-B:B-V* colour-colour diagram. For two of the clusters, we then see that the data in the *U-B:B-V* colour-colour diagram is not well fit by a Solar metallicity ZAMS, suggesting that metallicity is possibly more important in the estimation of Hubble's constant than previously thought.

9.2 Further Work Suggested by this Thesis

The work described in this Thesis has opened up many other areas that could be looked into either with resources used in this Thesis or with follow up observations and simulations. In this Section, we describe some of the projects that could extend the work presented here.

In Chapter 3 we applied a simple method of power spectrum analysis to the Durham /UKST Survey. This method was chosen to allow a comparison with other power spectra that were estimated in a similar way. However, Tegmark et al. (1998) describes other methods of power spectrum analysis, such as the Karhunen-Loève (K-L) method, that can be adjusted to allow for various systematic effects and biases, such as the integral constraint or obscuration by dust. The main advantage of this method is that the K-L coefficients retain spatial information, allowing processes which affect the radial and angular clustering differently to be optimally probed. For example, redshift space distortions only affect the clustering in the radial direction and can therefore be studied more cleanly through this way of quantifying the clustering. However, no direct comparison between the Fourier transform method and the K-L method of calculating the power spectrum from a large optically selected redshift survey has been made. This could be done using the Durham/UKST survey to check that on the scales where the FFT method of power spectrum analysis gives robust results the two methods of power spectrum analysis are equivalent.

As discussed above and in Chapter 4, more information on galaxy clustering can be obtained if higher order clustering statistics are measured. As well as using the method of counts-in-cells to estimate higher order clustering, N-point correlation functions can also be measured. The benefit of the N-point correlation functions over counts-in-cells is that information on the geometrical distribution of galaxies is preserved. However, even with only 2,500 galaxies (and corresponding random points) in the Durham/UKST survey, for $N > 2$ this is still computationally intensive to calculate. Methods for speeding up the estimation of the N-point correlation function are required which can be tested on the smaller surveys in advance of the completion of the 2dF GRS and SDSS which will contain at least 250,000 galaxies.

The main aim of Chapter 5 was to construct mock catalogues that had similar clustering to that expected in the 2dF QSO Survey in order to test clustering estimators and errors. However, we were only able to test the estimators under the assumption that

Ω_m=0.3 and Ω_Λ=0.7 was the correct cosmology. We would like to repeat the tests with lightcone simulations with different cosmologies, e.g. simulations run with an Einstein-de Sitter or an open cosmology. In particular, we would then be able to check that the Λ test results are independent of the cosmology of the simulation and that discrimination between different cosmologies will be possible what ever the cosmology of the Universe.

In order to relate the clustering in the 2dF QSO Survey to the clustering of the mass, the QSO-mass bias must be understood. However, the model used to bias the dark matter particles in the *Hubble Volume* was not physically motivated. Other methods of biasing the dark matter particles are described in the literature and could be applied to a simulation such as the *Hubble Volume*, although as the particle mass in the *Hubble Volume* is $2.2\times10^{12}M_\odot$ we may require a higher resolution simulation to do this. The models of Matarrese et al. (1997) predict the clustering bias of halos of a given mass, based on the work of Mo & White (1996). We can test if QSOs live in the most massive halos or if they live in a range of halo mass, by comparing the clustering statistics of the simulation with those measured from the 2dF QSO Survey. Similarly, Colín et al. (1999) identify halos in a simulation with a given circular velocity. The clustering at $z = 0$ matches the clustering seen on the APM correlation function (Baugh 1996) if halos that host APM galaxies have a circular velocity larger than $\sim$120 km s^{-1}. The clustering of halos with circular velocities larger than V_{max} as a function of redshift could be compared with the QSO clustering at high redshift, using a simulation with a lightcone output. Semi-analytic models of QSO formation are also starting to be developed (Kauffmann & Haehnelt 2000). By combining these models with N-body simulations the clustering of the models can be compared to QSO clustering. The sensitivity of the results to the different assumptions made in the model can also be tested in this way.

There is also further work required in our understanding of the Cepheid Period-Luminosity relation. The most important result from our work on the Galactic Open clusters is that the *U-B:B-V* data for two of the clusters does not match the ZAMS line. We suggest that this may be due to effects of metallicity as line blanketing is known to redden the colours of F and G type stars in metal rich clusters, causing the F and G stars to have a UV deficit in the colour-colour diagram. Metal poor F and G stars, on the other hand, have a UV excess. One of the best ways to test this would be to obtain intermediate resolution spectroscopy of a sample of F and G stars in this cluster to look at the strength of metal lines such as CaII, H, K, FeI and CaI. If metallicity is affecting the Galactic cluster Cepheids to the extent that is seen here, then there will be a significant

effect on Hubble's constant.

9.3 The Future

The current generation of redshift surveys have provided a wealth of information on the distribution of objects in the Universe and have clearly demonstrated the importance of redshift surveys in the area of cosmology. However, the results are currently limited by factors such as the number of objects in the survey, the depth of the survey, and the area covered by the survey. There are, however, a number of surveys that are currently under way or that are planned to start in the not too distant future that should increase our knowledge of large scale structure.

As mentioned in many of the Chapters, the future of wide angle galaxy surveys at $z \sim 0$ lies with the 2dF GRS (see for example Colless 1998) and the Sloan Digital Sky Survey (SDSS) (Gunn & Weinberg 1995). These surveys will be able to measure the correlation function and power spectrum out to large scales with high accuracy, providing very strong constraints on the present day distribution of galaxies which must be matched by models of structure formation. The wider angle of the galaxy survey may possibly permit measurements of the power spectrum out to scales which overlap with results from the COBE experiment, allowing the galaxy-mass bias to be measured directly. The surveys will also contain enough galaxies that they can be divided up as a function of luminosity, spectral type or environment, providing yet more constraints on the models of structure formation as well as galaxy formation. The value of the small scale peculiar velocities of galaxies and the parameter $\beta_{\rm gal}$ will be measured to high accuracy and any variation of these parameters as a function of scale may be detected. Other statistics, such as counts-in-cells and topological quantities, will also be measured from these surveys, providing a very detailed description of the distribution of local galaxies.

The 2dF GRS and SDSS surveys will only measure the redshifts of galaxies in the local Universe. There are two planned surveys that will study galaxies with an average redshift of $z \sim 1$. These are the DEEP Survey (Davis & Faber 1998) and the VIRMOS Survey (Le Fèvre 1998). These surveys will allow the galaxy clustering and galaxy luminosity function to be measured as a function of redshift, which will enable the clustering of QSOs and galaxies to be compared at higher redshifts than is currently possible. QSOs seem to have a similar clustering amplitude to present day galaxies, as shown in this Thesis, and to $z \sim 3$ Lyman break galaxies (Adelberger et al. 1998). However, we do not know how

strongly galaxies at $z \sim 1-2$ cluster so the relationship between galaxies and QSOs over this range of redshifts is not known. Our knowledge of galaxy clustering at even higher redshifts will be possible with the extension of the technique of Steidel et al. (1999) to higher redshifts. Galaxies that are detected in the V band and redder wavebands but not in the U or B band, known as B band drop outs, are expected to have a redshift of $z \sim 4$. Considerable numbers (~ 250) of these objects have already been detected. There is even tentative evidence that this method can be extended to detect galaxies at $z \sim 12$ using J band drop outs (Dickinson 2000).

As with the wide angle galaxy surveys, the future of QSO surveys lies with the completion of the 2dF and the SDSS QSO surveys. Once these are completed then the clustering of optically selected QSOs will be measured from QSOs with $z < 3$. All the different measurements of the clustering statistics applied to the mock 2dF QSO catalogues, such as the correlation function, power spectrum and Λ test, will have to be applied to the completed 2dF QSO Survey. Only then will we really know whether or not cosmology can be constrained by the 2dF QSO Survey. There are many more statistics that perhaps could be applied to the 2dF QSO Survey that have not been considered here, such as counts-in-cells, the three-point correlation function, topological quantities etc, although the relatively low space density of QSOs may limit some of the results. It will be interesting to see if there is consistency between higher order statistics, such as skewness, measured from galaxy and QSO surveys.

As discussed in Chapter 5, the 2dF QSO Survey will contain enough QSOs that clustering as a function of redshift can be studied. This means that clustering as a function of other properties, such as QSO colours and magnitudes, can also be studied. This may help in our understanding of QSO physics if certain classes of QSOs are found to cluster more strongly than others.

The surveys described above are either underway and or will be started very soon. In the more distant future, a new telescope, VISTA, will come online which will carry out wide angle, photometric surveys of galaxies and QSOs to deeper magnitude limits than have previously been possible. VISTA will also observe galaxies and QSOs in many different wavebands, allowing photometric redshifts to be calculated. Follow up spectroscopic work may be possible on the AAT if the AAOmega instrument, which is a proposed extension to the 2dF instrument, gets the go ahead. This could extend both the 2dF GRS and QSO Survey to wider angles and greater depths allowing, for example, the present day galaxy clustering to be measured even more accurately and the evolution of QSO

clustering to be studied in more redshift bins out to even higher redshifts.

Most of the work described in this Thesis has been based on optically selected surveys. Surveys can also be extended to other wavebands. In Chapter 2, we discuss various far-infrared surveys. Currently the Two-Micron All Sky Survey (2MASS) is being carried out which will allow the clustering in the near-infrared to be measured. With the launch of the Chandra and Newton satellites, redshifts of X-ray selected galaxies and QSOs may also be measured and there are techniques that allow redshifts in radio wavebands to be determined. In Chapter 3, we have shown that the clustering amplitude of far-infrared selected galaxies is lower than that of optically selected galaxies. Objects selected in X-ray and radio wavebands may have a different clustering amplitude again. Knowledge of the amplitude of galaxy clustering as a function of wavelength may yield important results for models of galaxy formation.

Over the last 20-30 years, galaxy and QSO surveys have revolutionised our understanding of large scale structure. This trend is set to continue with ambitious projects, that are currently underway or that are in the pipeline. These surveys will yield ever more accurate and powerful constraints on cosmological parameters and models of structure formation and will further increase our knowledge of the structure and scale of the Universe.

Appendix A | *Cosmology and Redshift*

In Chapter 1, the standard model and the Cosmological Principle were discussed. If the Cosmological Principle is assumed, then the Universe is described by a special metric known as the Friedmann-Robertson Walker (FRW) metric,

$$d\tau^2 = c^2dt^2 - a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2)\right]. \tag{A.1}$$

Here r, θ and ϕ are comoving spherical polar coordinates, t is proper time, τ is the space time interval and k is known as the curvature constant and determines the geometry of the Universe. The function $a(t)$ is the expansion factor of the Universe and is determined by solving the Einstein field equations. The solutions, know as the Friedman equations, are

$$\frac{1}{a}\ddot{a} = \frac{4}{3}\pi G\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda c^2}{3} \tag{A.2}$$

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8}{3}\pi G\rho - \frac{kc^2}{a^2} + \frac{\Lambda}{3} \tag{A.3}$$

where Λ is the cosmological constant, which arises as a constant of integration in solving the Einstein field equations and has dimensions of length^{-2}. p is the pressure, ρ is the density, both due to the matter and radiation in the Universe, and G is the gravitational constant. The dot represents a derivative with respect to cosmological proper time t.

The critical density of a flat Λ=0 Universe, is given by

$$\rho_c = \frac{3}{8\pi G}\left(\frac{\dot{a}}{a}\right)^2 \tag{A.4}$$

and this allows the important quantity, Ω_m to be defined as

$$\Omega_m = \frac{\rho}{\rho_c}. \tag{A.5}$$

Ω_Λ is defined as

$$\Omega_\Lambda = \frac{\Lambda c^2}{3H^2}. \tag{A.6}$$

Rearranging equation A.3 and applying equation A.4, we can derive

$$\Omega_m + \Omega_\Lambda = 1 + \frac{kc^2}{H^2 a^2}, \tag{A.7}$$

where H is the Hubble parameter, defined as $H(t) = \dot{a}(t)/a(t)$. Again, if we assume Λ=0, then if k=0 then the Universe is flat and the rate of expansion of the Universe will converge to zero. If $k > 0$ then the Universe will at some point stop expanding and start to collapse back onto itself, possibly resulting in the 'Big Crunch', the opposite of the 'Big Bang'. If $k < 0$ then the Universe is open and will expand forever. Observational evidence, such as the latest results analysis of the Cosmic Microwave Background (Lange et al. 2000), seem to suggest $\Omega = \Omega_m + \frac{\Lambda c^2}{3H^2} \sim 1$ and results from Supernova experiments (Perlmutter et al. 1999) suggest $\Lambda > 0$.

A.1 Relation between Redshift and Expansion Factor

Consider photons emitted at times t_e and $t_e + \delta t_e$ from a galaxy at comoving distance r away. For a photon $d\tau$=0 so for a radially moving photon integrating equation A.1 gives

$$\int_{t_e}^{t_0} \frac{cdt}{a(t)} = \int_0^r \frac{dr}{\sqrt{1-kr^2}} = \int_{t_e+\delta t_e}^{t_0+\delta t_0} \frac{cdt}{a(t)} \tag{A.8}$$

For small δt (i.e. $\delta t << H(t)$) we can rewrite this as

$$0 = \int_{t_e}^{t_0} \frac{cdt}{a(t)} + \left(\frac{c}{a(t_0)}\delta t_0 - \frac{c}{a(t_e)}\delta t_e\right) - \int_{t_e}^{t_0} \frac{cdt}{a(t)} \tag{A.9}$$

The first and last terms cancel leaving

$$\frac{\delta t_0}{a(t_0)} = \frac{\delta t_e}{a(t_e)} \tag{A.10}$$

If we now identify δt_e and δt_0 with the period of a light wave then $\delta t_e = 1/\nu = \lambda_e/c$ and similar for δt_0, where ν and λ are the frequency and wavelength of the wave. Therefore

$$\frac{\lambda_0}{a(t_0)} = \frac{\lambda_e}{a(t_e)}. \tag{A.11}$$

and by definition

$$1 + z = \frac{\lambda_0}{\lambda_e} = \frac{a(t_0)}{a(t_e)}. \tag{A.12}$$

A.2 Distances from Redshift

The FRW metric, equation A.1, allows the distance of a galaxy to be determined from its redshift.

Again, for a photon, $d\tau$=0. If the photon is moving in a purely radial direction so that $d\theta$ and $d\phi = 0$ then

$$cdt = \pm a(t)\frac{dr}{\sqrt{1-kr^2}} \tag{A.13}$$

by integrating from t_e to the present day we find

$$\int_{t_e}^{t_0}\frac{cdt}{a(t)} = -\int_r^0 \frac{dr}{\sqrt{1-kr^2}} \tag{A.14}$$

where t_e is the time at which the photon is emitted from the galaxy and t_0 is the time at which the photon is received by the observer. The negative sign is normally taken as we consider a photon which is moving towards us.

Now replacing $dt = \frac{da}{\dot{a}}$

$$\int_{a_e}^{a_o}\frac{cda}{a(t)\dot{a}} = -\int_r^0 \frac{dr}{\sqrt{1-kr^2}} = \begin{cases} \sin(r) & \text{if } k>0 \\ r & \text{if } k=0 \\ \sinh(r) & \text{if } k<0 \end{cases} \tag{A.15}$$

where $a_e = a(t_e)$ and $a_o = a(t_0)$. Let us call this f(a_e).

Now consider the comoving distance to a galaxy at the present day. $dt = 0$ in this case so

$$\int(-d\tau^2)^{1/2} = \int_0^r a_0\frac{dr}{\sqrt{1-kr^2}} = a_0 f(a_e). \tag{A.16}$$

Therefore the comoving distance, $D(z)$, is given by

$$D(z) = a_0\int_{a_e}^{a_0}\frac{cda}{a^2H(a)} \tag{A.17}$$

where Hubble's parameter ($= \dot{a}/a$) has been introduced. Finally the comoving distance can be written in its usual form, replacing a_0/a by $1+z$, equation (A.12),

$$D(z) = \int_0^z \frac{cdz}{H(z)}. \tag{A.18}$$

Hubble's parameter is dependent on the cosmology. Substituting the expressions for ρ_{crit} and Ω_Λ at present day into equation A.3, we obtain

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{H_o^2\rho}{\rho_{\text{crit}}} + \frac{kc^2}{a^2} + \Omega_\Lambda H_o^2. \tag{A.19}$$

Multiply the top and bottom of the ρ_{crit} term by ρ_o and replace ρ_{crit} and use the fact that $\rho \propto a^{-3}$ to obtain

$$\left(\frac{\dot{a}}{a}\right)^2 = H_o^2\left(\frac{a_o}{a}\right)^3\Omega_m + \frac{kc^2}{a^2} + \Omega_\Lambda H_o^2. \tag{A.20}$$

We now need to consider the term kc^2/a^2. At present day, we can rewrite equation A.3 as

$$\left(\frac{\dot{a}}{a_o}\right)^2 = H_o^2 \left(\frac{a_o}{a_o}\right)^3 \Omega_m + \frac{kc^2}{a_o^2} + \Omega_\Lambda H_o^2 \tag{A.21}$$

therefore

$$kc^2 = a_o^2 H_o^2 (1 - \Omega_m - \Omega_\Lambda) \tag{A.22}$$

and equation A.20 becomes

$$\left(\frac{\dot{a}}{a}\right)^2 = H_o^2 \left(\frac{a_o}{a}\right)^3 \Omega_m + H_o^2 (1 - \Omega_m - \Omega_\Lambda) \left(\frac{a_o}{a}\right)^2 + \Omega_\Lambda H_o^2 \tag{A.23}$$

and as $a \propto (1+z)^{-1}$ we finally obtain

$$H(z)^2 = H_0^2 \left\{(1 - \Omega)(1+z)^2 + \Omega_\Lambda + \Omega_m (1+z)^3\right\} \tag{A.24}$$

with $\Omega = \Omega_m + \Omega_\Lambda$.

If $\Omega = \Omega_m = 1$ then equation A.18 simplifies to

$$D(z) = \frac{2c}{H_0}\left(1 - \frac{1}{\sqrt{1+z}}\right) \tag{A.25}$$

Alternatively if $\Omega = \Omega_m + \Omega_\Lambda = 1$ then equation A.18 simplifies to

$$D(z) = \frac{c}{H_0} \int_0^z \frac{dz'}{\sqrt{(1 - \Omega_m) + \Omega_m (1+z')^3}} \tag{A.26}$$

Appendix B | *Linear Theory*

B.1 Linear Theory

In Newtonian theory, the evolution of a self gravitating fluid is governed by three equations. These are the continuity (mass conservation) equation, Euler's equation and Poisson's equation:

$$\frac{\partial \rho}{\partial t} + \nabla.\rho\mathbf{v} = 0 \tag{B.1}$$

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}.\nabla)\mathbf{v} + \frac{1}{\rho}\nabla p + \nabla\varphi = 0 \tag{B.2}$$

$$\nabla^2\varphi - 4\pi G\rho = 0 \tag{B.3}$$

where ρ is the density, p is the pressure and $\mathbf{v}$ is the velocity of the fluid, φ is the gravitational potential. ∇ denotes differentiation with respect to the position $\mathbf{r}$. If any dissipative terms arising from viscosity or thermal conductivity are ignored, then the conservation of entropy is also required:

$$\frac{\partial s}{\partial t} + \mathbf{v}.\nabla s = 0. \tag{B.4}$$

A static solution, with $\rho = \rho_0$, $p = p_0$, $\mathbf{v} = 0$ and $\nabla\varphi = 0$ is permitted. However, if ρ_0 is not equal to zero, the gravitational potential must vary spatially and so the Universe will either expand or contract. Hence a static universe is unstable.

However, the universe is not static but is expanding. The expansion of the Universe acts to slow down the rate at which fluctuations grow as the fluctuation essentially finds it harder to attract nearby material which is moving away from it.

The equations above apply for physical coordinates, but to solve the equations in an expanding universe scenario, it is more convenient to express the equations in terms of comoving position, $\mathbf{x} = \mathbf{r}/a$, a conformal time, $\mathrm{d}\tau = \mathrm{d}t/a$, and a peculiar velocity, $\mathbf{V} = a\mathrm{d}\mathbf{x}/\mathrm{d}t = \mathbf{v}$ - $\mathbf{x}\mathrm{d}a/\mathrm{d}t$ (note that $\mathbf{x}\mathrm{d}a/\mathrm{d}t = (\mathbf{x}/\mathrm{a})\mathrm{d}a/\mathrm{d}\tau$ which is the velocity due to the expansion of the Universe). It can then be shown (see, for example, Coles and Lucchin 1995) that the three equations of motion can then be rewritten as

$$\dot{\delta} + \nabla[(1+\delta)\mathbf{V}] = 0 \tag{B.5}$$

$$\dot{\mathbf{V}} + \mathbf{V}.\nabla\mathbf{V} + \mathbf{V}\dot{a}/a = -\nabla p/\bar{\rho} - \nabla\varphi \tag{B.6}$$

$$\nabla^2\varphi = 4\pi G\bar{\rho}a^2\delta \tag{B.7}$$

where $\bar{\rho}$ is the mean density of the Universe, $\delta = (\rho - \bar{\rho})/\bar{\rho}$ is the overdensity, ∇ is now differentiation with respect to $\mathbf{x}$ and $\dot{}$ is differentiation with respect to τ.

From these three equations, one can obtain the expression

$$\ddot{\delta} + \frac{\dot{a}}{a}\dot{\delta} = \frac{\nabla p}{\bar{\rho}} - \nabla^2\varphi. \tag{B.8}$$

From this equation, the growth of perturbations in different models can be understood. If we assume a dust model, for which the pressure is zero, and that the Universe is static, in which case there is no $\dot{a}/a$ term, then the overdensities, δ, have a growing and a decaying mode which are both exponential functions of time, with e-folding time $\sqrt{4\pi G\bar{\rho}}$. However, we have already stated that a static universe is unstable.

If we again assume a dust model but now in an expanding universe, then, in an Einstein-de Sitter cosmology, $a \propto \tau^2$ and the overdensities grow in proportion to $t^{2/3}$.

If the universe is radiation dominated but we still assume an Einstein-de Sitter universe, then the pressure term is $p = 1/3\rho c^2$ and the overdensities can be show to grow in proportion to t. For other cosmologies see Peebles (1980, 1984).

B.2 Relation of the Velocity Field to the Density Field

The equations of motion have so far been used to determine the rate of growth of fluctuations. However, the equations of motion also determine the velocity field $\mathbf{v}$ and the gravitational potential φ.

If only the first order terms are considered in the above expressions, equation B.5 reduces to

$$\dot{\delta} + \nabla.\mathbf{V} = 0 \tag{B.9}$$

as we assume $\delta \ll 1$. As the velocity field is irrotational, $\mathbf{V}$ can be expressed as the gradient of some velocity potential Φ_v so that

$$\nabla.\mathbf{V} = -\nabla^2\Phi_v = -\dot{\delta} \tag{B.10}$$

This can be rewritten as

$$\nabla^2\Phi_v = \delta\frac{\dot{\delta}}{\delta}\frac{a}{\dot{a}}H = fH\delta \tag{B.11}$$

where $f = (a\dot{\delta})/(\dot{a}\delta)$. This has the approximate form of $f \sim \Omega_{\mathrm{m}}^{0.6}$ which is the term that appears in expressions of redshift space distortions.

Using Poisson's equation in the form of equation B.7, the velocity field and the gravitational potential can be related by

$$\nabla^2\Phi_v = \nabla^2\varphi\frac{fH}{4\pi G\bar{\rho}}, \tag{B.12}$$

therefore the velocities are related to the overdensity though

$$\nabla.\mathbf{V} = -fH\delta. \tag{B.13}$$

B.3 The Power Spectrum

The overdensity, $\delta(\mathbf{x})$, was earlier defined as $(\rho(\mathbf{x}) - \bar{\rho})/\bar{\rho}$. This can be expressed as a Fourier series

$$\delta(\mathbf{x}) = \sum_{\mathbf{k}} \delta_{\mathbf{k}}\exp(i\mathbf{k}.\mathbf{x}) = \sum_{\mathbf{k}} \delta_{\mathbf{k}}^{*}\exp(-i\mathbf{k}.\mathbf{x}). \tag{B.14}$$

The wavenumber $\mathbf{k}$ has components

$$k_i = n_i\frac{2\pi}{L} \tag{B.15}$$

where i represents each of the three components of the vector $\mathbf{k}$, n_i are integers and L is the length of a cubical volume V. The Fourier coefficients, $\delta_{\mathbf{k}}$, are complex quantities given by

$$\delta_{\mathbf{k}} = \frac{1}{V}\int_V \delta(\mathbf{x})\exp(-i\mathbf{k}.\mathbf{x})d\mathbf{x}. \tag{B.16}$$

This power spectrum is then defined as $P(k) =< |\delta_{\mathbf{k}}|^2 >$, where the average is taken over all modes with $|\mathbf{k}|=k$.

Bibliography

The following abbreviations are used in this bibliography:

A&A: Astronomy and Astrophysics

AJ: The Astronomical Journal

ApJ: The Astrophysical Journal

ApJL: The Astrophysical Journal Letters

ApJS: The Astrophysical Journal Supplement

ARA&A: Annual Reviews of Astronomy and Astrophysics

JRASC: Journal of the Royal Astronomical Society of Canada

MNRAS: Monthly Notices of the Royal Astronomical Society

PASP: Publication of the Astronomical Society of the Pacific

Adelberger K. L., Steidel C. C., Givalisco M., Dickinson M., Pettini M. & Kellogg M., 1998, ApJ, 505, 18

Alcock C. & Paczyński B., 1979, Nature, 281, 358

Allen C. W., Astrophysical Quantities, 3rd Edition, 1973, The Athlone Press

Allen P. & Shanks T., 2000, in preparation

Andreani P. & Cristiani S., 1992, ApJ, 398, 13

Arp H., 1958, ApJ, 128, 166

Arp H., 1960, ApJ, 131,321

Arp H., Sandage A. & Stephens C., 1959, ApJ, 130, 80

Bahcall N. A., 1977, ARA&A, 15, 505

Balbi A. et al., 2000, astro-ph/0005124

Ballinger W. E., Peacock J. A. & Heavens A. F., 1996, MNRAS, 282, 877

Bardeen J. M., Bond J. R., Kaiser N. & Szalay A., 1986, ApJ, 304, 15

Barrell S., 1980, ApJ, 240, 145

Baugh, C.M., 1996, MNRAS, 280, 267.

Baugh C.M. et al., 2000, in preparation

Baugh C.M. & Efstathiou G., 1993, MNRAS, 265, 145

Baugh C.M. & Efstathiou G., 1994, MNRAS, 267, 323

Baugh C.M., Gaztañaga E. & Efstathiou G., 1995, MNRAS, 274, 1049.

Bean A. J., Efstathiou G., Ellis R. S., Peterson B. A. & Shanks T., 1983, MNRAS, 205, 605

Becker W. & Fenkert R., 1971, A&AS, 4, 241

Benoist C., Cappi A., Da Costa L. N., Mauogordato S., Bouchet F. R. & Schaeffer R., 1999, ApJ, 514, 563.

Benson A. J., Baugh C.M., Cole S. M., Frenk C. S. & Lacey C. G., 2000a, astro-ph/9910488

Benson A. J., Cole S. M., Frenk C. S., Baugh, C. M. & Lacey C. G., 2000b, MNRAS, 311, 793

Bouchet F. R., Schaeffer R. & Davis M., 1991, ApJ, 383, 19.

Bouchet F. R., Strauss M. A., Davis M., Fisher K. B., Yahil A. & Huchra J. P., 1993, ApJ, 417, 36.

Bower R. G., Coles P., Frenk C. S. & White S. D. M., 1993, ApJ, 405, 403

Boyle B. J., 1986, PhD Thesis, Univ of Durham

Boyle B. J., Jones L. R. & Shanks T., 1991, MNRAS, 251, 482

Boyle B. J., Shanks T., Croom S. M., Smith R. J., Miller L., Loaring N. & Heymans C., 2000, astro-ph/0005368

Cameron L. M., 1984, A&A 146, 59

Cameron L. M., 1985, A&A 147, 39

Carney B. W., 1979, ApJ, 233, 211

Carroll S. M., Press W. H. & Turner E. L., 1992, ARA&A, 30, 499

Cole S. M., Fisher K. B. & Weinberg D. H., 1995, MNRAS, 267, 785

Cole S. M., Hatton S. J., Weinberg D. H. & Frenk C. S., 1998, MNRAS, 300, 945

Coles P., 1993, MNRAS, 262, 1065.

Coles P. & Lucchin F., 1995, "Cosmology", John Wiley & Sons LtD, Chichester

Colín P., Klypin A. A., Kravtsov A. V. & Khokhlov A. M., 1999, ApJ, 523, 32

Colless M., 1998, 'Looking Deep in the Southern Sky', eds Morganti R. & Couch W. J., ESO/Australia Workshop, Springer, Pg 9

Collins C. A., Heydon-Dumbledon N. H. & MacGillivray H. T., 1988, MNRAS, 236, 7p

Colombi S., Bouchet F.R. & Hernquist L., 1996, ApJ, 465, 14.

Colombi S., Szapudi I. & Szalay A. S., 1998, MNRAS, 296, 253.

Coulson I. M. & Caldwell J. A. R., 1984, SAAO Preprint, 337

Crampton D., Schade D. & Cowley A. P., 1985, AJ, 90, 987

Cristiani S. et al., 1995, A&AS, 112, 347

Cristiani S., 1999, in Proceedings of the VLT Opening Symposium, Antofagasta, Chile, astro-ph/9905314

Croft R. A. C., Weinberg D. H., Katz N. & Hernquist L., 1998, ApJ, 495, 44

Croom S. M., et al, 2000, in preparation

Croom S. M., Ratcliffe A., Parker Q. A., Shanks T., Boyle B. J. & Smith R. J., 1998, MNRAS, 306, 592

Croom S. M. & Shanks T., 1996, MNRAS, 281, 893

Croom S. M., Shanks T., Boyle B. J., Smith R. J., Miller L. & Loaring N., 1998, Evolution of Large Scale Structure: From Recombination to Garching

Davis M., Efstathiou G., Frenk C. S. & White S. D. M., 1985, ApJ, 292, 371

Davis M. & Faber S. M., 1998, in 'Wide Field Surveys in Cosmology', Editions Frontieres, ISBN 2-8 6332-241-9, 333

de Bernardis P. et al., 2000, Nature, 404, 955

de Lapparent V., Kurtz M. J. & Geller M. J., 1986, ApJ, 304, 585

de Vaucouleurs G., 1948, Ann. d'Astrophysique, 11, 247

de Vaucouleurs G. & de Vaucouleurs A., 1964, Reference Catalogue of Bright Galaxies, University of Texas Press, IC1964

Dickinson M., 1999, in 'Physical Transactions of the Royal Society Series A', astro-ph/0004319

Efstathiou G., Bond J.R. & White S. D. M., 1992, MNRAS, 258, 1

Efstathiou G., Kaiser N., Saunders W., Lawrence A., Rowan-Robinson M., Ellis R. S. & Frenk C. S., 1990, MNRAS, 247, 10p

Efstathiou G., Sutherland W. J. & Maddox S. J., 1990, Nature, 348, 705

Eggen O. J., 1983, AJ, 88, 379

Eisenstein D. J. & Zaldarriaga M., 1999, astro-ph/9912149

Eke V. R., Cole S. M. & Frenk C. S., 1996, MNRAS, 282, 263

Ellingson E., Yee H. K. C. & Green R.F., 1991, ApJ, 371, 49

Evrard A. et al., 2000, in preparation.

Faber S. M. & Gallagher J. S., 1979, ARA&A, 17, 135

Feast M. W., 1957, MNRAS, 117, 193

Feast M. W. & Catchpole R. M., 1997, MNRAS, 286, L1

Feast M. W. & Walker A. R., 1987, ARA&A, 25, 345

Feldman H. A., Kaiser N. & Peacock J. A., 1994, ApJ, 426, 23

Ferrarese L. et al., 2000, ApJ, 529, 745

Fisher K. B., Huchra J. P., Strauss M. A., Davis M., Yahil A. & Schlegel D., 1995, ApJS, 100, 69

Frieman J. A. & Gaztañaga E., 1994, ApJ, 425, 392

Fry N. J., 1996, ApJ, 461, L65

Fry A. & Carney B., 1997, AJ, 113, 1073

Fry J. N. & Gaztañaga E., 1993, ApJ, 413, 447

Fry J. N. & Gaztañaga, E., 1994, ApJ, 425, 1

Gaztañaga E., 1992, ApJ, 398, L17

Gaztañaga E., 1994, MNRAS, 268, 913

Gaztañaga E., 1995, ApJ, 454, 561

Gaztañaga E. & Baugh C. M., 1995, MNRAS, 273, L1

Gaztañaga E. & Baugh C. M., 1998, MNRAS, 294, 229

Gaztañaga E. & Bernardeau F., 1998, A&A, 331, 829

Gaztañaga E., Fosalba P. & Elizalde E., 1999, astro-ph/9906296

Gaztañaga E. & Frieman J. A., 1994, ApJ, 437, L13

Geisler D., 1987, AJ, 94, 84

Georgantopoulos I. & Shanks T., 1994, MNRAS, 271, 773

Gibson B. K. et al., 2000, ApJ, 529, 723

Gould A., 1994, ApJ, 426, 542

Graham J. A., 1982, PASP, 94, 244

Gratton R. G., Fusi Pecci F., Carretta E., Clementini G., Corsi C. E. & Lattanzi M., 1997, ApJ, 491, 749

Gregory S. A. & Thompson L. A., 1978, ApJ, 222, 784

Guetter H. H., 1991, AJ, 103, 197

Gunn J. E. & Weinberg D. H., 1995, in proceedings of the 35th Herstmonceux workshop, Cambridge University Press, Cambridge

Guzzo L., 1999, 'Large Scale Structure at the Turn of the Millennium', in Proceedings of the XIX Texas Symposium on Relativistic Astrophysics, Paris, December 1998

Guzzo L. et al., 2000, A&A, 355, 1

Hale-Sutton D., 1990, PhD Thesis, University of Durham

Haiman Z. & Hui L., 2000, astro-ph/0002190

Hamilton A. J. S., 1992, ApJ, 385, L5

Hamilton A. J. S., 1993, ApJ, 417, 19

Hamilton A. J. S., 1998, in 'The Evolving Universe', ed D. Hamilton, p185, Kluwer Academic

Hanany S. et al., 2000, astro-ph/0005123

Hatton S. J., 1999, PhD Thesis, University of Durham

Hivon E., Bouchet F. R., Colombi S. & Juszkiewicz R., 1995, A&A, 298, 643

Hewitt P. C., Foltz C. B. & Chaffee F. H., 1995, AJ, 109, 1498

Hoag A. A., Johnson H. L., Iriarte B., Mitchell R. I., Hallam K. L. & Sharpless S., 1961, Publ. US. Nav. Obs, 17, 347

Hubble E., 1929, Proc. Nat. Acc. Sci., 15, 168

Huchra J. P., Davis M., Latham D. & Tonry J., 1983, ApJS, 72, 433

Hui L. & Gaztañaga E., 1999, ApJ, 519, 622

Humason M. L., Mayall N. U. & Sandage A. R., 1956, AJ, 61, 97

Hunt L. K., Mannucci F., Testi L., Migliorini S., Stanga R. M., Baffa C., Lisi F. & Vanzi L., 1998, AJ, 115, 2594

Janes K. A., 1979, ApJS, 39, 135

Jenkins A. et al., 1998, ApJ, 499, 20

Johnson H. L., 1960, ApJ, 131, 620

Juszkiewicz R., Bouchet F. & Colombi S., 1993, ApJ, 412, L9

Kaiser N., 1987, MNRAS, 227, 1

Kauffmann G., Colberg J. M., Diaferio A. & White S. D. M., 1999, MNRAS, 303, 188.

Kauffmann G. & Haehnelt M., 2000, MNRAS, 311, 576

Kennicutt R. et al., ApJ, 498, 181

Kibblewhite E. J., Bridgeland M. T., Bunclark P. & Irwin M. J., 1984, Astronomical Microdensitometry Conference, NASA Conf. Publ., 2317, 277

Kim R. S. J. & Strauss M. A., 1998, ApJ, 493, 39

Kirschner R. P., Oemler A. & Schechter P. L., 1978, AJ, 83, 1549

Koo D. C., Kron R. G. & Cudworth K. M., 1986, PASP, 98, 285

Kraft R. P., 1958, ApJ, 128, 1610

Krisciunas K., Sinton W., Tholen K., Tokunaga A., Golisch W., Griep D., Kaminski C., Impey C. & Christian C., 1987, PASP, 99, 887

Landolt A. U., 1992, AJ, 104, 372

Landy S. D. & Szalay A. S., 1993, ApJ, 412, 64

Laney C. D. & Stobie R. S., 1993, MNRAS, 263, 921

Laney C. D. & Stobie R. S., 1994, MNRAS, 266, 441

Lange A. E. et al., 2000, astro-ph/0005004

Lawrence A., Rowan-Robinson M., Ellis R.S., Frenk C. S., Efstathiou G., Kaiser N., Saunders W., Parry I. R., Xiaoyang X. & Crawford J., 1999, MNRAS, 308, 897

Lawrence C. R. & Lange A. E, 1997, AAS, 191, 4120

Le Fèvre O. et al., 1998, in 'Wide Field Surveys in Cosmology', Editions Frontieres, ISBN 2-8 6332-241-9, 333

Lin H., Kirschner P., Shectman S. A., Landy S. D., Oemler A., Tucker D. L. & Schechter P. L., 1996, ApJ, 471, 617

Ling E. N., Frenk C. S. & Barrow J. D., 1986, MNRAS, 223, 21P

Loveday J., Maddox S. J., Efstathiou G. & Peterson B. A., 1995, ApJ, 442, 457

Loveday J., Peterson B. A., Efstathiou G. & Maddox S. J., 1992, ApJ, 390, 338

Loveday J., Peterson B. A., Maddox S. J. & Efstathiou G., 1996, ApJS, 107, 201

McCracken H. J., 1999, PhD Thesis, University of Durham

McCracken H. J., Metcalfe N. & Shanks T., 2000, in preparation

MacGillivray H. T. & Stobie R. S., 1984, Vistas Astr., 27, 433

Maddox S. J., Efstathiou G. & Sutherland W. J., 1996, MNRAS, 283, 1227

Maddox S. J., Sutherland W. J., Efstathiou G. & Loveday J., 1990, MNRAS, 243, 692

Madore B. F., 1975, A&A, 38, 471

Madore B. F. & Van den Bergh S., 1975, ApJ, 197, 55

Mann R. G., Peacock J. A. & Heavens A. F., 1998, MNRAS, 293, 209.

Martini P. & Weinberg D., 2000, astro-ph/0002384

Matarrese S., Coles P., Lucchin F. & Moscardini L., 1997, MNRAS, 286, 115

Mather J. C. et al., 1994, ApJ, 420, 439

Matsubara T. & Suto Y., 1996, ApJ, 470, L1

Menzies J. W., Marang F., Laing J. D., Coulson I. M. & Engelbrecht C. A., 1991, MNRAS, 248, 642

Mermilliod J. C., 1981, A&A, 97, 235

Mészáros P., 1974, A&A, 37, 225

Miller L., et al., 2000, in preparation

Mo H. J., Jing Y. P. & Börner G., 1992, ApJ, 392, 452

Mo H. J. & White S. D. M., 1996, MNRAS, 282, 347

Moffat A. F. J. & Vogt N., 1975, A&AS, 20, 155

Moscardini L., Coles P., Lucchin F. & Matarrese S., 1998, MNRAS, 299, 95

Narayanan V., Weinberg D. H., Branchini E., Frenk C. S., Maddox S., Oliver S., Rowan-Robinson M. & Saunders W., 1999, astro-ph/9910229

Osmer P. S., 1981, ApJ, 247, 762

Panagia N. & Tossi M., 1980, A&AS, 81, 375

Park C., Vogeley M. S., Geller M. J. & Huchra J. P., 1994, ApJ, 431, 569

Parker Q. A. & Watson F. G., 1995 in Maddox S. J. & Aragón-Salamanca A., eds., 35th Herstmonceux Conf. Cambridge, Wide Field Spectroscopy and the Distant Universe, Princeton Univ. Press, Princeton, NJ, p33

Peacock J. A., 1991, MNRAS, 253, 1

Peacock J. A., 1997, MNRAS, 284, 885

Peacock J. A. & Dodds S. J., 1994, MNRAS, 267, 1020

Peacock J. A. & Dodds S. J., 1996, MNRAS, 280, L19

Peacock J. A. & Nicholson D., 1991, MNRAS, 253, 307

Pedreros H., Madore B. F. & Freedman W. L., 1984, ApJ, 286, 563

Peebles P. J. E., 1980. "The Large Scale Structure in the Universe", Princeton University Press, Princeton

Peebles P. J. E., 1984, ApJ, 284, 439

Penzias A. A. & Wilson R. W., 1965, ApJ, 142, 421

Percival W. & Miller L., 1999, MNRAS, 309, 823

Perlmutter A. et al., 1999, ApJ, 517, 565

Peterson B. A., Ellis R. S., Efstathiou G., Shanks T., Bean A. J., Fong R. & Zen-Long Z., 1986, 221, 233

Popowski P. A., Weinberg D. H., Ryden B. S. & Osmer P. S., 1998, ApJ, 498, 11

Press W. H. & Schechter P., 1974, ApJ, 187, 425

Rana N. C., 1991, ARA&A, 29, 129

Ratcliffe A., 1996, PhD Thesis, University of Durham

Ratcliffe A., Shanks T., Parker Q. A. & Fong R., 1996, MNRAS, 281, L47

Ratcliffe A., Shanks T., Parker Q. A. & Fong R., 1998a, MNRAS, 293, 197

Ratcliffe A., Shanks T., Parker Q. A. & Fong R., 1998b, MNRAS, 296, 197

Ratcliffe A., Shanks T., Parker Q. A. & Fong R., 1998c, MNRAS, 296, 191

Ratcliffe A., Shanks T., Parker Q. A., Broadbent A., Watson F. G., Collins C. A. & Fong, R., 1998d, MNRAS, 300, 417

Renzini A. et al., 1996, ApJL, 459, L23

Romeo G., Bonifazi A., Fusi Pecci F. & Tosi M., 1989, MNRAS, 240,459

Saha A., Sandage A., Tammann G. A., Labhardt L., Macchetto F. D. & Panagia N., 1999, ApJ, 522, 802

Sandage A., 1958, ApJ,128, 150

Sandage A., 1960, ApJ, 131, 610

Sandage A. & Tammann G. A., 1981, A Revised Shapley-Ames Catalogue of Bright Galaxies, Washington DC, Carnegie Institute of Washington, RSA

Sandage A., Veron P. & Wyndham J. D., 1965, 142, 1307

Saunders W., Frenk C. S., Rowan-Robinson M., Lawrence A. & Efstathiou G., 1991, Nature, 349, 32

Saunders W., Sutherland W. J., Maddox S. J., Keeble O., Oliver S. J., Rowan-Robinson M., McMahon R. G., Efstathiou G. P., Tadros H., White S. D. M., Frenk C. S., Carraminana A. & Hawkins M. R. S., 2000, astro-ph/0001117

Shectman S. A., Landy S. D., Oemler A., Tucker D. L., Lin H., Kirschner R. P. & Schechter P. L., 1996, ApJ, 470, 172

Schmidt B. P. et al., 1998, ApJ, 507, 46

Schmidt M., 1963, Nature, 197, 1040

Schmidt M. & Green R. F., 1983, ApJ, 269, 352

Schmidt-Kaler T., 1982, in Shaifers K. & Voigt H. H., eds, Landolt-Bornstein VI 2b, Springer-Verlag, Berlin, p. 12

Schramm N. D., 1991, Physica Scripta, T36, 22

Seljak U. & Zaldarriaga M., 1996, ApJ, 469, 437

Shane C. D. & Wirtanen C. A., 1967, Pub. Lick. Obs., vol 22, Part 1

Shanks T., Fong R. & Boyle B. J., 1983, Nature, 303, 156

Shanks T., Fong R., Boyle B. J. & Peterson B. A., 1986, In IAU Symp No 119 on Quasars, eds Swarup G. & Kapahi V. K., Reidel, Dordrecht, Holland

Shanks T., Fong R., Boyle B. J. & Peterson B. A., 1987, MNRAS, 271, 753

Shanks T. & Boyle B. J., 1994, MNRAS, 271, 753

Sharpless S., 1963, in Basic Astronomical Data, ed. K. A. Strand, University of Chicago Press,, p. 225

Shaver P. A., 1984, A&A, 136, L9

Shu F. H., The Physical Universe, University Science Books, 1982, 286

Smith R. J., Boyle B. J. & Maddox S. J., 1995, MNRAS, 277, 270

Smith R. J., Boyle B. J. & Maddox S. J., 2000, MNRAS, 313, 252

Smoot G. F. et al., 1992, ApJL, 396, L1

Steidel C. C., Adelberger K. L., Givalisco M., Dickinson M. & Pettini M., 1999, ApJ, 519, 1

Steidel C. C., Pettini M. & Hamilton D., 1995, ApJ, 110, 2519

Strauss M. A. & Willick J. A., 1995, Physics Reports, 261, 271

Sugiyama N., 1995, ApJS, 100, 281

Sutherland W., Tadros H., Efstathiou G., Frenk C. S., Keeble O., Maddox S., McMahon R. G., Oliver S., Rowan-Robinson M., Saunders W. & White S. D. M., 1999, MNRAS, 308, 289

Szapudi I. et al., 2000, in preparation

Szapudi I., Colombi S. & Bernardeau F., 1999a, MNRAS, 310, 428

Szapudi I. & Gaztañaga E., MNRAS, 1998, 300, 493.

Szapudi I., Meiksin A. & Nichol R. C., 1996, ApJ, 473, 15

Szapudi I., Quinn T., Stadel J. & Lake G., 1999b, ApJ, 517, 54

Szapudi I. & Szalay A. S., 1993, ApJ, 408, 43.

Tadros H. & Efstathiou G., 1995, MNRAS, 276, L45

Tadros H. & Efstathiou G., 1996, MNRAS, 282, 1381

Tadros H., Efstathiou G. & Dalton G., 1998, MNRAS, 296, 995

Talbert F. D., 1975, PASP, 87, 341

Tanvir N. R., Ferguson H. C. & Shanks T., 1999, MNRAS, 310, 175

Tanvir N. R., Shanks T., Ferguson H. C. & Robinson D. R. T., 1995, Nature, 377, 27

Tegmark M., Hamilton A., Strauss M., Vogeley M. & Szalay A., 1998, ApJ, 499, 555

Thackeray A. D., Wesselink A. J., & Harding G. A., 1962, MNRAS, 124, 445

Tokunaga A. T., 1998, in Astrophysical Quantities (Revised), 4th edition, submitted - see UKIRT web page.

Turner D. G., 1979a, JRASC, 73, 74

Turner D. G., 1979b, PASP, 91, 642

Turner D. G., 1980, ApJ, 240, 137

Turner D. G., 1981, AJ, 86, 231

Turner D. G., 1984, PASP, 96, 422

Turner D. G., 1993a, ApJS, 85, 119

Turner D. G., 1993b, A&AS, 97, 755

Turner D. G., Forbes D. & Pedreros M., 1992, AJ, 104, 3, 1132

Turner D. G., Pedreros M. H. & Walker A. R., 1998, AJ, 1115, 1958

Tytler D., Fan X. M. & Burles S., 1996, Nature, 381, 207

van den Bergh S., 1978, AJ, 83

van den Bergh S. & Harris G. L. H., 1976, ApJ, 208, 765

van Leeuwen F., 1999, A&A, 341, 71

Veron-Cetty M. P. & Veron P., 1984, in A Catalogue of Quasars and AGN, ESO Scientific Report Series, 1

Veron-Cetty M. P. & Veron P., 1998, in A Catalogue of Quasars and AGN, ESO Scientific Report Series, 18

Walker A. R., 1985a, MNRAS, 213, 889

Walker A. R., 1985b, MNRAS, 214, 45

Walker A. R. & Laney C. D., 1986, MNRAS, 224, 61

Welch D. L., McAlary C. W., Madore B. F., McLaren R. A. & Neugebauer G., 1985, ApJ, 292, 217

White S. D. M., Efstathiou G. & Frenk C. S., 1993, MNRAS, 262, 1023

White S. D. M., Frenk C. S. & Davis M., 1983, ApJ, 274, 1

White M., Gelmini G. & Silk J., 1995, Physical Review D, 51, 2669

Wright E. L., 1999, New Astronomy Reviews, 43, 257

Yee H. K. C. & Green R. F., 1987, ApJ, 319, 28

Yee H. K. C., Morris S. L., Lin H., Carlberg R. G., Hall P. B., Sawicki M., Patton D. R., Wirth G. D., Ellingson E. & Shepherd C. W., 2000, astro-ph/0004026

Zaritsky D., Kennicutt R. C. & Huchra J. P., 1994, ApJ, 420, 87

Zwicky F., 1933, Helvetica Phys. Acta., 6, 110